

PR 18-SEP-1997; 97US-0059263.
 PR 18-SEP-1997; 97US-0059266.
 PR 15-OCT-1997; 97US-0062125.
 PR 17-OCT-1997; 97US-0063285.
 PR 17-OCT-1997; 97US-0063287.
 PR 21-OCT-1997; 97US-0063486.
 PR 24-OCT-1997; 97US-0062814.
 PR 24-OCT-1997; 97US-0062816.
 PR 24-OCT-1997; 97US-0063045.
 PR 24-OCT-1997; 97US-0063120.
 PR 24-OCT-1997; 97US-0063121.
 PR 24-OCT-1997; 97US-0063122.
 PR 27-OCT-1997; 97US-0063128.
 PR 27-OCT-1997; 97US-0063329.
 PR 28-OCT-1997; 97US-0063541.
 PR 28-OCT-1997; 97US-0063542.
 PR 28-OCT-1997; 97US-0063544.
 PR 28-OCT-1997; 97US-0063549.
 PR 28-OCT-1997; 97US-0063550.
 PR 29-OCT-1997; 97US-0063564.
 PR 29-OCT-1997; 97US-0063435.
 PR 29-OCT-1997; 97US-0063704.
 PR 29-OCT-1997; 97US-0063732.
 PR 29-OCT-1997; 97US-0063738.
 PR 29-OCT-1997; 97US-0063739.
 PR 29-OCT-1997; 97US-0064215.
 PR 31-OCT-1997; 97US-0063870.
 PR 31-OCT-1997; 97US-0064103.
 PR 03-NOV-1997; 97US-0064248.
 PR 07-NOV-1997; 97US-0064809.
 PR 12-NOV-1997; 97US-0065186.
 PR 17-NOV-1997; 97US-0065846.
 PR 18-NOV-1997; 97US-0065693.
 PR 21-NOV-1997; 97US-0066120.
 PR 21-NOV-1997; 97US-0066364.
 PR 24-NOV-1997; 97US-0066772.
 PR 24-NOV-1997; 97US-0066466.
 PR 24-NOV-1997; 97US-0066770.
 PR 24-NOV-1997; 97US-0066511.
 PR 24-NOV-1997; 97US-0066453.
 (GETH) GENENTECH, INC.
 Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
 WPI: 1999-229533/15.
 N-PSDB; AAX52231.
 New isolated human genes and polypeptides used in, e.g. treatment of
 gastrointestinal ulceration
 Claim 12; Fig 36; 520pp; English.
 AAX13344-403 represent secreted and transmembrane human proteins.
 The cDNA sequences are obtained from cDNA libraries, prepared from
 fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
 The encoded polypeptides have specific uses based on their homology to
 known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
 associated with the preservation and maintenance of gastrointestinal
 mucosa and the repair of acute and chronic mucosal lesions
 (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
 ulceration and congenital microvillus atrophy), skin diseases associated
 with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
 cancers such as lung squamous cell carcinoma, of the vulva and gliomas),
 potent effects on cell growth and development, diseases related to
 growth or survival of nerve cells including Parkinson's disease,
 Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
 a target for anti-tumor drugs. PRO533 may be used in the treatment
 of Usher Syndrome or Atrophila areata. PRO269 can be used as an
 anti-thrombotic agent; PRO287 polypeptides and portions may have

CC therapeutic applications in wound healing and tissue repair; PRO317 can
 be used for treating problems of the kidney, uterus, endometrium, blood
 vessels, or related tissue, e.g. in the heart or genital tract.
 XX Sequence 490 AA;
 SO
 Query Match 100.0%; Score 2605; DB 20; Length 490;
 Best Local Similarity 100.0%; Pred. No. 8, 2e-200;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPAFALCLMOLMPEPGGGEHPTADRACSSAGCYSLHATMNRKOAEEACILRGCA 60
 DB 1 mrfafalclmqlmpgpgggehptadragccsagcyslhathmrkrgaeecillrgca 60
 QY 61 LSTVRAGAEIRAVALLRAGPGGGSKDLLFVVALEERRSHTLNEPLRGFSWLSDDP 120
 DB 61 lstrvragaelravallragpgggskdllfvaleterrshcleneplrgfswlsddp 120
 QY 121 GGLSDTLQWVEEPORSCARRCAVLQATGVGPAGMKEMRCHLRANGYLCKYQFEVLCF 180
 DB 121 gglstdtlqwveeporscarrcaavlqatgvgepagmkemrchlranngylckyqfevlcf 180
 QY 181 APPRGAASNSTYRAPHOLHSALDESPGETEVYALCRGOLPISTVTCIADETIGARMDKLSG 240
 DB 181 apprgaasnstyrappholhsaldespgetevyalcrgolpistvtciadetigarmdkls 240
 QY 241 DVLCPGGRYLRAGKAEPLNCLDLDGFPACBACATGEELGKGRSGVNSGEGOPTLGIG 300
 DB 241 dvlcpgrylragkaeplnclldlgfpacbacatgeelgkgrsgvnsgegoptlgig 300
 QY 301 VPTRRPATATSPVQRTWPIRYDEKLTPLVPEQDMSVTSIPETPMWSOSTMSTLQM 360
 DB 301 vptrrpatatspvqrtwpirydekltplvpeqdmsvtsipetpmwsostmstlqm 360
 QY 361 SLQAEKATITPSSGVYSKFNSTSSAPFOADSSAVFIFRVAVVVLIITFTVGL 420
 DB 361 slqaekatitpssgvyskfnstssapfoadssavfifrvavvvliitftvgl 420
 QY 421 VLCTHESPSQPKRESMGPGLGSDPEPALGSSSAHCTNNGKVVSGCDLDRDAEGALL 480
 DB 421 vlcthespsqpkresmgpplgdsdpelpalgsasahctnngkvvgcdldrdaegall 480
 QY 481 AESPLGSSDA 490
 DB 481 aesplgssda 490
 RESULT 2
 AAB33460
 ID AAB33460 standard; Protein: 490 AA.
 AC AAB33460;
 XX
 DT 29-JAN-2001 (first entry)
 XX
 DE Human. PRO269 protein UNQ236 SEQ ID NO:233.
 XX
 KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;
 KW dermatological; antiarthritic; antirheumatic; immunosuppressive;
 KW hemostatic; antitumor; antidiabetic; neurotropic; neuroprotective;
 KW antianemic; hepatotropic; virucide; antipsoriatic; antiallergic;
 KW antidiabetic; systemic lupus erythematosus; rheumatoid arthritis;
 KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
 KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
 KW autoimmune thrombocytopenia; immune mediated renal disease;
 KW demyelinating disease; hepatobiliary disease; Whipple's disease;
 KW inflammatory bowel disease; gluten-sensitive enteropathy;
 KW autoimmune disease; immune-mediated skin disease; allergic disease;
 KW immunological disease; transplantation associated disease;
 KW graft rejection; graft-versus-host-disease.
 XX

OS Homo sapiens.
 XX WO200053758-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05841.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99US-0123618.
 PR 12-MAR-1999; 99US-0123957.
 PR 23-MAR-1999; 99US-0125775.
 PR 12-APR-1999; 99US-0128849.
 PR 20-APR-1999; 99WO-US08615.
 PR 28-APR-1999; 99US-0131445.
 PR 04-MAY-1999; 99US-0132371.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 06-JUL-1999; 99US-0145698.
 PR 08-JUL-1999; 99US-0146222.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30999.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;
 XX
 DR WPI: 2000-572271/53.
 DR N-PSDB; AAC58625.
 XX
 PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 XX
 PS Claim 33; Fig 94; 309pp; English.
 XX
 CC The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central

CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
 CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
 CC autoimmune or immune-mediated skin diseases, allergic diseases,
 CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host-disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.
 XX
 SQ Sequence 490 AA:
 Query Match 100.0%; Score 2605; DB 21; Length 490;
 Best Local Similarity 100.0%; Pred. No. 8.2e-200;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPAPALCLIMQALMPGPGGGEHPTDRAGCSAGCYSLHARTMKROAAEEACILRGGA 60
 DB 1 mRPAPALCLIMQALMPGPGGGEHPTDRAGCSAGCYSLHARTMKROAAEEACILRGGA 60
 QY 61 LSTVRAGAEALRAVLALIRAGPGGSGSKDLFWALERRRSHCTLENEPLRGFSWLSDP 120
 DB 61 lSTVRAGAEALRAVLALIRAGPGGSGSKDLFWALERRRSHCTLENEPLRGFSWLSDP 120
 QY 121 GGLIESDTLQWVEEPQSRCTARRCAVLQATGVPEAGWKEMRCHLRANGYLCKYQFEVLCP 180
 DB 121 gGLIESDTLQWVEEPQSRCTARRCAVLQATGVPEAGWKEMRCHLRANGYLCKYQFEVLCP 180
 QY 181 ARPRGASNLSTRAPQLHSAALDFSPRTEVSALCRGQLPISVYTIADIEIGARMKLSG 240
 DB 181 aRPrgaSnLStrApQLhSaALDFSPrTEvSALCRGQLPISVYTIADIEIGARMKLSg 240
 QY 241 DVLPCPGRYLRAGKCAEPLNCLIDLGCFACACATGFEIGKDGKSCVTSGEQPTLGATG 300
 DB 241 dVLpcpGRyLRAGKCAePLNCLIDlgCFACAcATGfEIGKDGKScVTSgEQpTLGATg 300
 QY 301 VPTRRPAPATSPVPORTWPIRYDEKLGFTPLVPEODNSVTSIPELPRNGSOSTMSTLQM 360
 DB 301 vPTRRpAPATSPVORTWPIrYDEKLGfTPLVPEODNSVtSIPElPRNGSOSTMstLqM 360
 QY 361 SIQAESKATITPSSGVSISFENSTSSATPOAPSSAAVFIPIFSTAVVVLITLTVLGL 420
 DB 361 SIqAESKATITPSSGVSISfENSTSSATpOAPSSAAVFIPIfSTAVVVLITLTVlGL 420
 QY 421 VKLCFHESSPSQPRKESMGPGLIESDPEPAALGSSSAHCTNNGVKGDCLDRARAGAL 480
 DB 421 vKLCfHESSpSQPRKESmGPGLIESDPEpAALGSSSAhCTNNGVKGDCLDRARAGaL 480
 QY 481 AESPLGSSDA 490
 DB 481 aEsPlGSSda 490
 RESULT 3
 AAB24403 standard; Protein; 490 AA.
 AAB24403;
 07-NOV-2000 (first entry)
 Human PRO269 protein sequence SEQ ID NO:77.
 Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
 diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
 angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;
 cytotatic; gene therapy; vaccine.
 Homo sapiens.
 WO200032221-A2.

PD 08-JUN-2000.
 XX 30-NOV-1999; 99WO-US28313.
 XX 01-DEC-1998; 98WO-US25108.
 PR 16-DEC-1998; 98US-0112850.
 PR 12-JAN-1999; 99US-0115554.
 PR 08-MAR-1999; 99WO-US05028.
 PR 12-MAR-1999; 99US-0123957.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20594.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 XX (GENE) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A,
 PI Godowski PJ, Gurney AL, Klein RD, Kuon SS, Paoni NF, Smith V,
 PI Watanabe CK, Williams PM, Wood WI,
 DR WPI: 2000-412154/35.
 DR N-PSDB; AAA77570.
 XX Nucleic acids encoding PRO polypeptides useful for preventing,
 PT diagnosing and treating disorders in mammals
 PT angiogenic disorders in mammals
 XX Claim 72, Fig 32, 315pp; English.

XX The present invention describes nucleic acids encoding PRO polypeptides
 CC useful for preventing, diagnosing and treating disorders in mammals
 CC cardiovascular, endothelial or angiogenic disorder in mammals by
 CC modulating cell proliferation, angiogenesis and cardiovascularisation,
 CC and for identifying agonists and antagonists of these processes. The
 CC nucleic acids and the proteins they encode may be used in the
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate PRO expression such as cardiovascular, endothelial or
 CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
 CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
 CC containing them and the PRO polypeptide may be used to treat disorders
 CC associated with decreased PRO expression. AAA77510 to AAA77721 and
 CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
 CC the exemplification of the present invention.
 XX Sequence 490 AA:

Query Match 100.0%; Score 2605; DB 21; Length 490;
 Best Local Similarity 100.0%; Pred. No. 8; 2e-200;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPALCLIMQALMPGGGHPTRADRGCSAGACSLHMTWKROAEACTLRGA 60
 DB 1 mpatalclimqalwpgpggghpctradrgcsagacyshmtwkrgaeactlr 60
 QY 61 LSTVRGAELRAVLALLRAGPGGSKDLFWALERRHCHTENPLRGFMSLSDP 120
 DB 61 lstrvrgaelravalllrpgpggskdlflwalerrhchclenprrgfgmslssdp 120
 QY 121 GLESDTLQWMEERORSTARCAVLAQGVGEPAGWEMKCHDRANGYLCKYFVLCP 180
 DB 121 glesdltlwmeerorstarcavlaqgvgepagwemkchdrangylckyfevclcp 180
 QY 181 APRGASNLSTYRAPFOLHSAALDFSPGCTEVSALCRGOLPISTCTIADLGARWDLKLG 240

DB 181 aprgasnlsyrapfplhlaaldfspgctevsalcrgolpistctiadlgarwcklsg 240
 QY 241 DYLCPCGRYLRAKCAELPNCIDLDLGFACCEGATGELKDRSCVTSGEGPTLGATG 300
 DB 241 dylpcpgrylragkcaelpnciddlgfaccgatgkdrscvtsgsgptlgatg 300
 QY 301 VPTRRPATATSPVPOATWPIRYDEKLGRLPLVPEQDNSTSTLPELRNGSOSTMSTLQW 360
 DB 301 vptrrpatatspvpqatwpiirvdekigrplvpeqdnststlpeirngsostmstlqm 360
 QY 361 SLOAESKATTPSGSVLSKENTSSATPOAFDSSAAVFIPTAVVVLITMTVLGL 420
 DB 361 slgaeskattpgsvlskntssatpofdssavvfiptavvvllmtvlg 420
 QY 421 VKLCFHSPSSQPRKESKMPGLESPEPAALGSSAHCNTNNGVKGDCDLDRAGALL 480
 DB 421 vkcfhspssqprkesmvglespepaalgssahcntngvkgdcdldraegall 480
 QY 481 AESPLGSSDA 490
 DB 481 aesplgssda 490

RESULT 4
 ID AAY95016
 ID AAY95016 standard; Protein; 490 AA.

XX AAY95016;

DT 19-JUN-2000 (first entry)

DE Human secreted protein vp15_1, SEQ ID NO:72.

KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 KW neurodegenerative disease; asthma; contraceptive.

OS Homo sapiens.

PM WO200011015-A1.

PD 02-MAR-2000.

PF 24-AUG-1999; 99WO-US19351.

PR 24-AUG-1998; 98US-0097638.

PR 09-SEP-1998; 98US-0097659.

PR 28-SEP-1998; 98US-0102092.

PR 25-NOV-1998; 98US-0109978.

PR 23-DEC-1998; 98US-0113645.

PR 23-DEC-1998; 98US-0113646.

PR 23-AUG-1999; 99US-0379246.

PA (ALPH-) ALPHAGENE INC.

PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

DR WPI: 2000-224657/19.

PT New secreted or transmembrane proteins and polynucleotides encoding

PT them, useful for treating neurodegenerative disorders, autoimmune

PT diseases and cancer

PS Claim 81, Page 331-332; 357pp; English.

CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC and cDNA sequences encoding them (AAA23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:

CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis; regulation; tissue growth activity;
 CC activity/inhibin activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. The present sequence represents one of the
 CC 40 proteins of the invention.

Sequence 490 AA:

Query Match 100.0%; Score 2605; DB 21; Length 490;
 Best Local Similarity 100.0%; Pred. No. 8.2e-200;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPAFALCLMLQALMPGSGEHPADRACSGAGACSLHHMTKROAEACILRGGA 60
 DB 1 mrfafalcilmlqalmpgsggehpdracsgagacslhmtkrqaeacilrgga 60
 QY 61 LSTVRAGAEIRAVLALLRAGPGGGSKDLLFWVALERRSHCTLENEPLRGFSWLSDD 120
 DB 61 lstrvragaelravallrpgpggskdllfwvalerrshctleneplrgfswlsddp 120
 QY 121 GGESEDTLQWVEERORSCTARRCAVQATGTVBPAGKMKRCHLRANGYLCCKQFEVLC 180
 DB 121 ggeesetlqwveerpqscarrcaavqatgtvbpagkkmrchlranngylckkyfevlcp 180
 QY 181 APRGASNLSTYRAPFQLSHALDFSPGTEVVALCRGQPISTVCADTIGARWDLTSG 240
 DB 181 aprgasaanlstyrappqlshaldfspgtevalcrqgpistvctadigardwklsg 240
 QY 241 DVLCPCGRYLRAGKCAELPNCIDLDLGFACGECATGFEIAGKDRSCVTSGEQPTLG 300
 DB 241 dvlpcpgrylragkcaelpncidldlgfaccatgfeigkdrscvtsgsqptlqg 300
 QY 301 VPRRRPRTATSPVPORTWPIRVDKLGEPVLPBQDNSTSTIPETPRMGSSQSTMTLQM 360
 DB 301 vprrrprrtatsvpportwpirvdklgepvlpbqdnststipetprmgssqstmtlqm 360
 QY 361 SLQAEKATITPGSVYSIKNSTSTSPATPAFDSSAVVIFPSTAVVILMTMTVLGL 420
 DB 361 slqaeekattitpgsvysiknststspatpafdsavvifpstavviltmtvlgl 420
 QY 421 WKICFHESSSQPRKESMGPGLESDEPAALGSSAHCTNNGVKWDCCDLRRAEGLAL 480
 DB 421 wkicfhesssqprkesmgpglesdepaalgsahctnngvkwddclrraegall 480
 QY 481 AESSPLGSSDA 490
 DB 481 aessplyssda 490

RESULT 5

AAU12349 standard: Protein; 490 AA.

AAU12349;

24-OCT-2001 (first entry)

XX

DE Human PRO269 polypeptide sequence.
 XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIa; gene therapy.

OS Homo sapiens.

PN MO200140466-A2.

XX 07-JUN-2001.

PF 01-DEC-2000; 2000WO-US32678.

PR 01-DEC-1999; 99WO-US28301.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28554.

PR 02-DEC-1999; 99WO-US28555.

PR 09-DEC-1999; 99US-0170262.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 20-DEC-1999; 99WO-US30999.

PR 30-DEC-1999; 99WO-US31243.

PR 06-JAN-2000; 2000WO-US00277.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 18-FEB-2000; 2000WO-US04342.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.

PR 01-MAR-2000; 2000WO-US05601.

PR 20-MAR-2000; 2000WO-US07377.

PR 21-MAR-2000; 2000WO-US07532.

PR 30-MAR-2000; 2000WO-US08439.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.

PR 10-NOV-2000; 2000WO-US30873.

(GETH) GENENTECH INC.

XX Baker KP, Bersini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,

PI Gerritsen ME, Goddard A, Godowski PJ, Gueney AL, Sherwood S;

PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

DR WPI; 2001-408281/43.

DR N-PSDB; AAS21421.

PT Isolated, secretory and transmembrane PRO polypeptide used to detect

PT other PRO polypeptides, link bioactive molecules to cells expressing

PT PRO polypeptides, and detect the presence of mammalian tumours e.g.

PT lung, breast, prostate, cervical

XX Claim 12; Fig 356; 813p; English.

XX AAU12172-AAU12446 represent novel human secretory and transmembrane

XX PRO polypeptides. The PRO polypeptides are useful to detect other

XX PRO polypeptides, to link bioactive molecules to cells expressing

XX PRO polypeptides, to modulate biological activities of cells expressing

XX breast, prostate, rectal, cervical or liver tumours by comparing PRO

XX polypeptide expression in a cell sample to that in a control sample.

XX Some of the 275 sequences are also useful to stimulate the release of

XX tumour necrosis factor-alpha (TNF-alpha) from human blood, the

XX proliferation or differentiation of chondrocytes, the proliferation or

XX gene expression in pericyte cells, the release of proteoglycans from

XX cartilage, the proliferation of inner ear utricular supporting cells or

XX of T-lymphocytes, the release of a cytokine from peripheral blood

XX monocytes (PBMCs), or the proliferation of endothelial cells. Some of


```

Db 181 aprgaasnllyrapfqlhsaaldfspptevsalcrqqlpivtcladeigarwklsg 240
QY 241 DVLCPCGRYLRAKCAELPNCIDLDLGFAECATGFEELGKDRSCVTSGEQPTLGCTG 300
Db 241 dvlcpcgrylrakcaelpncidldlgfacecatgfeigkdrscvtsgeqptlgctg 300
QY 301 VPTRRPATATSPVPORTWPIRYDEKLGEPPLVPEODNSTVSTPEIPRMKSOSTMSTLQM 360
Db 301 vptrrpatatspvportwpiirvdekigepplvpeodnstvstpeiprmksostmstlqm 360
QY 361 SLOAESKATITPGSVYSKFNSTTSATPOAFDSSAVVIFSTAVVILTMVVLGL 420
Db 361 slgaeskatitpgsvyskfnsttsatpafdsasavvifstavvilltmvlgl 420
QY 421 VKLCFHESSOPRKESMGPPGLESDPEPALGSSSAHCTNNGVKQDCDLRRAEGALL 480
Db 421 vklcfhessoprkessmgppglesdpepalgssahctnngvkqdcclrdraegall 480
QY 481 AESPLGSSDA 490
Db 481 aesplgssda 490
DI 481 aesplgssda 490

RESULT 7
AAB80228
ID AAB80228 standard; Protein; 428 AA.
XX
AC AAB80228;
XX
DT 24-APR-2001 (first entry);
XX
DE Human PRO269 protein.
XX
KW Human; PRO; dermatological; antiproliferative; cytostatic; antiinflammatory;
KW antiparkinsonian neurotrophic; neuroprotective; vulnereary; cardiant;
KW antidiabetic; vasotrophic; antidiabetic; antirheumatic; cancer;
KW antihistaminic; antifertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.
XX
OS Homo sapiens.
XX
PN WO200104311-A1.
XX
PD 18-JAN-2001.
XX
PF 22-FEB-2000; 2000WO-US04414.
XX
PI 27-JUL-1999; 99US-0143048.
PI 28-JUL-1999; 99US-0145698.
PI 28-JUL-1999; 99US-0146222.
PI 08-SEP-1999; 99WO-US20594.
PI 13-SEP-1999; 99WO-US20944.
PI 15-SEP-1999; 99WO-US21090.
PI 15-SEP-1999; 99WO-US21547.
PI 05-OCT-1999; 99WO-US23089.
PI 29-NOV-1999; 99WO-US28214.
PI 30-NOV-1999; 99WO-US28313.
PI 16-DEC-1999; 99WO-US30095.
PI 20-DEC-1999; 99WO-US30911.
PI 20-DEC-1999; 99WO-US30999.
PI 05-JAN-2000; 99WO-US00219.
XX
PA (GENTH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gertschen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gunney AL, Hillan KJ, Kijavits ID;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PW, Wood WL;
XX
WPI: 2001-081051/09.
DR N-PSDB: AAF72389.

```

```

XX
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease)
XX
PS Claim 1; Fig 36; 393pp; English.
XX
CC The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosa.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.
XX
SQ Sequence 428 AA:

```

```

Query Match 85.6%; Score 2230; DB 22; Length 428;
Best Local Similarity 87.3%; Pred. No. 6; e-170;
Matches 428; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

```

```

QY 1 MRPAFALCILMOWLPBGSGGEPTADACSGACYSIHHMTMKROAEACILIRGA 60
Db 1 mrpafalcilwgalwpgsggeptadragcsagacsylhmtmkraeacilirga 60
QY 61 LSTVRGAELRAVLALIRAPGPGGSKDLFWVALERRSHCTLENEPLRGFSWLSDDP 120
Db 61 lstvrgaelrvallirapgggskdlfwvalerrshctlepeirgfwlsddp 120
QY 121 GGEESDTLOWVEBPQNSCTARCAVLOATGCVBPAGKEMKCHLRANGYLCKTQFEVLCP 180
Db 121 ggeesdtlowvebpqnsctarcavloatgcvbpagkcmchlrangylcktyfevlcp 180
QY 121 gglestdlqvwepqnsctarcavloatgcvbpagkcmchlrangylcktyfevlcp 180
QY 181 APRGAASNLISYRAPFQLHSAALDFSPGTEVSALCRGQLPISTVCIADIEIGARMDLSG 240
Db 181 aprga----- 186
QY 241 DVLCPCGRYLRAKCAELPNCIDLDLGFAECATGFEELGKDRSCVTSGEQPTLGCTG 300
Db 187 -----rylragkcaelpncidldlgfacecatgfeigkdrscvtsgeqptlgctg 238
QY 301 VPTRRPATATSPVPORTWPIRYDEKLGEPPLVPEODNSTVSTPEIPRMKSOSTMSTLQM 360
Db 239 vptrrpatatspvportwpiirvdekigepplvpeodnstvstpeiprmksostmstlqm 238
QY 361 SLOAESKATITPGSVYSKFNSTTSATPOAFDSSAVVIFSTAVVILTMVVLGL 420
Db 239 slgaeskatitpgsvyskfnsttsatpafdsasavvifstavvilltmvlgl 358
QY 421 VKLCFHESSOPRKESMGPPGLESDPEPALGSSSAHCTNNGVKQDCDLRRAEGALL 480
Db 359 vklcfhessoprkessmgppglesdpepalgssahctnngvkqdcclrdraegall 418
QY 481 AESPLGSSDA 490
Db 419 aesplgssda 428

RESULT 8
AA91410
ID AA91410 standard; Protein; 344 AA.
XX
AC AA91410;
XX
DT 29-JUN-2000 (first entry)
XX

```

Human secreted protein sequence encoded by gene 4 SEQ ID NO:131.

Human; secreted protein; diagnosis; neuroprotective; nootropic;
neuroleptic; anti-infective; cerebroprotective; immunomodulatory;
anti-microbial; cardiac; cytosolic; anti-inflammatory; haemostatic;
anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
neural; reproductive; immune disorder; immunodeficiency; infection;
lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
ischemia; mania; dementia; obsessive compulsive disorder;
viral prophylaxis; developmental disorder; sexually-linked disorder;
cardiovascular disorder; food additive; preservative; chromosome 14.

Homo sapiens.
WO200011014-A1.
02-MAR-2000.
24-AUG-1999; 99WO-US19330.
25-AUG-1998; 98US-0097917.
31-AUG-1998; 98US-0098634.
(HUMA-) HUMAN GENOME SCI INC.
Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA,
Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsu G,
Duan RD;
WPI: 2000-224656/19.

Novel secreted proteins and corresponding DNA molecules that can be
used to prevent, treat and diagnose disease in humans, for example,
Alzheimer's, cancer, and immune disorders

Disclosure: Page 391-392; 416pp; English

The polynucleotide sequences given in AAA26281 to AAA26336 encode the
human secreted proteins given in AAY91346 to AAY91449. The human secreted
proteins can have activities based on the tissues and cells they are
expressed in. Examples of the activities are: neuroprotective; nootropic;
neuroleptic; anti-infective; cerebroprotective; immunomodulatory;
anti-microbial; cardiac; cytosolic; anti-inflammatory; haemostatic;
anticonvulsant; and vasotropic. The polynucleotides and proteins may be
used to prevent, treat or ameliorate a medical condition, e.g. by protein
or gene therapy. Conditions treatable by the proteins of the invention
include neural, reproductive, or immune disorders, especially
immunodeficiency, infection, lymphomas, demyelinating diseases,
auto-immunities, cancer, general microbial infection, inflammation,
aneurysms and haemorrhages. Specific examples include: Alzheimer's
disease; Parkinson's; Huntington's; Tourette syndrome; multiple
sclerosis; meningitis; ischemia; prostate cancer; mania; dementia;
obsessive compulsive disorder and viral prophylaxis. The polynucleotides
and proteins can also be used in the detection of disorders associated
with the function of the protein, for example, the detection of
cardiovascular disorders, sexually-linked disorders, or disorders of the
cardiovascular system. They may also be used as food additives or
preservatives. AAA26272 to AAA26280 and AAY91345 are sequences used in
the exemplification of the present invention.

Sequence 344 AA:

Query Match 63.4%; Score 1652; DB 21; Length 344;
Best Local Similarity 98.1%; Pred. No. 7, 9e-124;
Matches 316; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

169 YLCKTQFVLCAPRPGASNTSYRPAQLHSAALDPSPGFVSALCRGQLPISVTGTA 228
23 YLCKYGFVLCAPRPGASNTSYRPAQLHSAALDPSPGFVSALCRGQLPISVTGTA 82

229 DEIGARMDKLSGDLVCPGGRYLRAGKCAELPNCIDLDGFRACCATGFEIGKDGRCVT 288
83 delgarwdklsgrdlcpgrgylragnkcaelnclldlgfraccatgfeigkdgrcvt 142
269 SGEQPTLCGCGVTRRPPATATSPVQRTPIVYDEKLGSETPLVPEQDMSVTISPEIRP 348
143 sgeqptlcgcvtrrppatatspvqrtpiyvdeklgetplvpeqdmsvtispeirp 202
349 WGSOSTMSTLOMSLOAESKATITPSSGVISKFNSTTSATPOAFDSSAVVFVSTAVY 408
203 wgsqstmsltmslqeskatitpssgviskfnsttsatpofadssavvfvfstavv 262
409 VLVITMTVIGVLCVCHESPSOPKESKGPPELESDEPPALGSSSAHCTNNGVKVG 468
263 vlvitmtvlgvlcchesspsopkeskgnppgvrwvllkpaalgssahctnnkvvgd 322
469 CDLRRAEGALLAESPLGSSDA 490
323 cdldraegallaesplgssda 344

RESULT 9

AAA41847
ID AAA41847 standard; Protein; 331 AA.

AAA41847;

08-FEB-2001 (first entry)

Human ORFX ORF1611 polypeptide sequence SEQ ID NO:3222.

Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
vulnerary; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; anti-inflammatory;
antiviral; antibacterial; antifungal; antipneumatic; antithyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive.

Homo sapiens.

WO200058473-A2.

05-OCT-2000.

31-MAR-2000; 2000WO-US08621.

31-MAR-1999; 99US-0127607.

02-APR-1999; 99US-0127636.

05-APR-1999; 99US-0127728.

30-MAR-2000; 2000US-0540763.

(CURA-) CURAGEN CORP.

Shinkets RA, Leach M;

WPI: 2000-602362/57.

N-PSDB; AAC76056.

Novel nucleic acids and peptides derived from open reading frame X,
useful for treating e.g. cancers, proliferative disorders,
neurodegenerative disorders and cardiovascular disease -

Claim 11; Page 2431-2432; 5507pp; English.

AAC74446 to AAC77606 encode the proteins given in AAA40237 to AAA43397;

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihydroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy.
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SC Sequence 331 AA:

Query Match 62.9%; Score 1639.5; DB 21; Length 331;
 Best Local Similarity 97.8%; Pred. No. 7.5e-123;
 Matches 310; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 128 LGWVEPQSCSTARRCAVQATGCVGPACMKERCHLRANGVYCKYQFELCPAPRGA 187
 DB 15 VQVVEPQSCSTARRCAVQATGCVGPACMKERCHLRANGVYCKYQFELCPAPRGA 73
 OY 188 SNUSTRAPFQLHSALDPSPTGVSAALCRGQLPSVTGIDAEIGARMKLSDVLPCP 247
 DB 74 SNUSTRAPFQLHSALDPSPTGVSAALCRGQLPSVTGIDAEIGARMKLSDVLPCP 133
 OY 248 GYTLRAGCAELPNCIDLDGFGACGATFELGKGRSCVTSSEGGPTLCGTCVPRRRP 307
 DB 134 GYTLRAGCAELPNCIDLDGFGACGATFELGKGRSCVTSSEGGPTLCGTCVPRRRP 193
 OY 308 AATATSPPORTPIRVDEKIGETPLVPEODNSVTSIPETPRMCSOSTMGLQMSLAQESK 367
 DB 194 AATATSPPORTPIRVDEKIGETPLVPEODNSVTSIPETPRMCSOSTMGLQMSLAQESK 253
 OY 368 AATTPSGSVISKFNSTTSATPOAEDSSAVVFIFVSTAVVVLITMTVLGLVKLFHE 427
 DB 254 AATTPSGSVISKFNSTTSATPOAEDSSAVVFIFVSTAVVVLITMTVLGLVKLFHE 313
 OY 428 SPSQPRKESMGPGLE 444
 DB 314 SPSQPRKESMGPGLE 330

RESULT 10

AAV91349 standard: Protein; 189 AA.

AAV91349;

29-JUN-2000 (first entry)

Human secreted protein sequence encoded by gene 4 SEQ ID NO:70.

Human; secreted protein; diagnosis; neuroprotective; nootropic;
 neuroleptic; antianemic; cerebroprotective; immunomodulatory;
 anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
 anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
 neural; reproductive; immune disorder; immunodeficiency; infection;
 lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
 aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
 Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
 ischaemia; mania; dementia; obsessive compulsive disorder;
 viral prophylaxis; developmental disorder; sexually-linked disorder;

KW cardiovascular disorder; food additive; preservative; chromosome 14.
 XX Homo sapiens.
 OS
 XX MO200011014-A1.
 PN-
 XX 02-MAR-2000.
 PD
 XX 24-AUG-1999; 99WO-US19330.
 PF
 XX 25-AUG-1998; 98US-0097917.
 PR 31-AUG-1998; 98US-0098634.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
 PI Soppet DR, Latleur DW, Endress GA, Ebner R, Komatsu G;
 PI Duan RD;
 DR WPI; 2000-224656/19.
 DR N-PSDB; AAA26284.
 XX
 PT Novel secreted proteins and corresponding DNA molecules that can be
 PT used to prevent, treat and diagnose disease in humans, for example,
 PT Alzheimer's, cancer, and immune disorders -
 XX
 PS Claim 11; Page 357; 416p; English.

CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the
 CC human secreted proteins given in AAV91346 to AAV91449. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: neuroprotective; nootropic;
 CC neuroleptic; antianemic; cerebroprotective; immunomodulatory;
 CC anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
 CC anticonvulsant; and vasotropic. The polynucleotides and proteins may be
 CC used to prevent, treat or ameliorate a medical condition, e.g. by protein
 CC or gene therapy. Conditions treatable by the proteins of the invention
 CC include neural, reproductive, or immune disorders, especially
 CC immunodeficiency, infection, lymphomas, demyelinating diseases,
 CC auto-immunities, cancer, general microbial infection, inflammation,
 CC aneurysms and haemorrhages. Specific examples include: Alzheimer's
 CC disease; Parkinson's; Huntington's; Tourette syndrome; multiple
 CC sclerosis; meningitis; ischaemia; prostate cancer; mania; dementia;
 CC obsessive compulsive disorder and viral prophylaxis. The polynucleotides
 CC and proteins can also be used in the detection of disorders associated
 CC with the function of the protein, for example, the detection of
 CC developmental disorders, sexually-linked disorders, or disorders of the
 CC cardiovascular system. They may also be used as food additives or
 CC preservatives. AAA26272 to AAA26280 and AAV91345 are sequences used in
 CC the exemplification of the present invention.
 XX

Sequence 189 AA:

Query Match 30.8%; Score 802.5; DB 21; Length 189;
 Best Local Similarity 84.4%; Pred. No. 2.9e-56;
 Matches 157; Conservative 0; Mismatches 26; Indels 3; Gaps 2;

OY 1 MRPAFLCLLMQALMPGPGGGEHTADRGCSASGACSLHHTMTKROAAEEACILRGGA 60
 DB 1 MRPAFLCLLMQALMPGPGGGEHTADRGCSASGACSLHHTMTKROAAEEACILRGGA 60
 OY 61 LSTVRAGAEIRAVLALLRAGPGPGGSKDLFWVALERRRSHCTLENEPLRGFSWLSDDP 120
 DB 61 LSTVRAGAEIRAVLALLRAGPGPGGSKDLFWVALERRRSHCTLENEPLRGFSWLSDDP 120
 OY 121 GGLSTTLQWVEPQSCSTARRCAVQATGCVGPACMKERCHLRANGVYCKYQFELCP 178
 DB 121 GGLSTTLQWVEPQSCSTARRCAVQATGCVGPACMKERCHLRANGVYCKYQFELCP 178
 OY 179 CVPAPR 184
 DB 180 rrapgp 185

RESULT 11
 AAY95031
 ID AAY95031 standard; Protein: 175 AA.
 AC AAY95031;
 DT 19-JUN-2000 (first entry)
 DE Human clone vpi5_1 insertional variant ORF. SEQ ID NO:131.
 KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
 blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 neurodegenerative disease; asthma; contraceptive; open reading frame;
 ORF; variant.
 OS Homo sapiens.
 WC2000011015-A1.
 PD 02-MAR-2000.
 PF 24-AUG-1999; 99WO-US19351.
 PR 24-AUG-1998; 98US-0097638.
 PR 24-AUG-1998; 98US-0097659.
 PR 09-SEP-1998; 98US-0099618.
 PR 28-SEP-1998; 98US-0102092.
 PR 25-NOV-1998; 98US-0109978.
 PR 23-DEC-1998; 98US-0113645.
 PR 23-DEC-1998; 98US-0113646.
 PR 23-AUG-1999; 98US-0379246.
 PA (ALPH-) ALPHAGEN INC.
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapley P;
 DR WPI: 2000-224657/19.
 XX New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer.
 PS Disclosure; Page 351; 357pp; English.
 CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC and CDNA sequences encoding them (AAA23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:
 CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activation/inhibition activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections; pulmonary inflammation; Guillain-Barre syndrome;
 CC rheumatoid arthritis; diabetes mellitus; and allergic reactions such as
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activating/inhibiting activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. Sequences AAY95024-Y95025, AAY95029 and
 CC AAY95031-Y95032 represent additional open reading frames (ORFs) that are
 CC encoded by deletion or insertional variants of the CDNA clones of the

CC invention.
 XX
 SO Sequence 175 AA;
 Query Match 29.08; Score 754.5; DB 21; Length 175;
 Best Local Similarity 94.08; P-Value 1.8e-52;
 Matches 142; Conservative 2; Mismatches 4; Indels 3; Gaps 1;
 QY 1 MRPAFALCLLMQALMPGGGHEPTADRACSGACYSIHHNPMKROAEACIIIRGA 60
 DB 1 mrapafalclllmqalwpgpggheptadragsasgacysihnmrqaeeacillrqa 60
 QY 61 LSTYRACAEIRAVIALIRAGPGGSGKLLFWALERRRSHCTLENEPLRGFSWLS 120
 DB 61 lstryracaelravialiragpgpggskllfwalerrrrshctleneplrgfswls 120
 QY 121 GLESDTLQWVEEPORSCTRRCACAVLCATGC 151
 DB 121 glesdtlqwveeporscstrrcacavlcattgc 148
 RESULT 12
 AAY91409
 ID AAY91409 standard; Protein: 123 AA.
 AC AAY91409;
 DT 29-JUN-2000 (first entry)
 DE Human secreted protein sequence encoded by gene 4 SEQ ID NO:130.
 KW Human; secreted protein; diagnosis; neuroprotective; neurotrophic;
 KW neuroleptic; antimanic; cerebroprotective; immunomodulatory;
 KW anti-microbial; cardiac; cytosolic; anti-inflammatory; haemostatic;
 KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
 KW neural; reproductive; immune disorder; immunodeficiency; infection;
 KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
 KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
 KW ischaemia; mania; dementia; obsessive compulsive disorder;
 KW viral prophylaxis; developmental disorder; sexually-linked disorder;
 KW cardiovascular disorder; food additive; preservative; chromosome 14.
 OS Homo sapiens.
 WC2000011014-A1.
 PD 02-MAR-2000.
 PF 24-AUG-1999; 99WO-US19330.
 PR 25-AUG-1998; 98US-0097917.
 PR 31-AUG-1998; 98US-0098634.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
 PI Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G;
 PI Dunn RD;
 DR WPI: 2000-224656/19.
 XX Novel secreted proteins and corresponding DNA molecules that can be
 PT used to prevent, treat and diagnose disease in humans, for example,
 PT Alzheimer's, cancer, and immune disorders.
 PS Disclosure; Page 391; 416pp; English.
 CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the
 CC human secreted proteins given in AAY91446 to AAY91449. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: neuroprotective; neurotrophic;

| | |
|-----------------|---|
| D | b |
| Oy | 261 ---NCLDLDLGGFAEACATGFE-----TLGGTGVPTRRRPPAR 309 |
| Bd | 414 gylvalgeedgqgdvdecwypggplcdslcftlgtsgflnoglpgvwlapngvstmgpyas 473 |
| Oy | 310 ATSPVPORTWPIKRD--EKIGETPLVPEDONSVTSLP-EIPRMGSOSTMTLOKSIOAES 366 |
| Bd | 474 lgpsspg---pdeekdykegest-vp---rataspitrgpegpkatptitsplsdsda 525 |
| Oy | 367 KAT-----LTPEGS-----VISFNFTSTSSAMPQAFDSSEA-----VFEL 401 |
| Bd | 526 pilsapikmlapsagsgswrepslhataasyggepagadsvsatgnmdtgdqklllfy 585 |
| Oy | 402 FVSTAVALVVILTMVLGLVKLCFHESSPSOPRKESMPCELSEDPPAA 451 |
| Bd | 586 ilglvtvaallillal-algl-lvykrkraareek-----ekkpnaa 625 |
| <hr/> | |
| RESULT_15 | |
| AAY32345- | ID AAVY32345 standard; Protein: 652 AA. |
| Xx | XX AC AAVY32345; |
| Xx | DJ 28-FEB-2000 (first entry) |
| Xx | DE Human cell surface receptor clqrp. |
| KM | ClaRp: receptor; cell surface; transmembrane; glycoprotein; human; host defence; infection; HIV; immunodeficiency; therapy; immunostimulant; phagocytosis; signal transduction. |
| OS | Homo sapiens. |
| FH | Key Location/Qualifiers Peptide 1..21 |
| FT | /note= "signal peptide" |
| FT | 22..652 |
| FT | /note= "mature protein" |
| FT | 22..580 |
| FT | /note= "extracellular domain" |
| FT | 581..605 |
| FT | /note= "transmembrane domain" |
| FT | 606..652 |
| FT | /note= "cytoplasmic domain" |
| FT | 260..301 |
| FT | /note= "epidermal growth factor-like domain 1" |
| FT | 302..344 |
| FT | /note= "epidermal growth factor-like domain 2" |
| FT | 345..384 |
| FT | /note= "epidermal growth factor-like domain 3" |
| FT | 385..426 |
| FT | /note= "epidermal growth factor-like domain 4" |
| FT | 427..468 |
| FT | /note= "epidermal growth factor-like domain 5" |
| FT | 325 |
| FT | /note= "N-glycosylated" |
| FT | 644 |
| Misc-difference | /note= "forms part of consensus motif recognised by tryosine kinases" |
| N | M09955839-AI. |
| DD | 04 -NOV-1999. |

XX 29-APR-1999; 99WO-US09335.
 PF 30-APR-1998; 98US-0071386.
 PR (REGC) UNIV CALIFORNIA.
 PA Tenner AJ, Nepomuceno RR;
 PI WPI; 2000-062021/05.
 DR N-PSDB; AA234989.
 XX
 XX A new cell surface receptor protein used as a prophylactic for
 PT individuals at risk from infection, e.g. HIV.
 PS
 PS Claim 2; Fig 3; 49pp; English.

This sequence represents a novel human cell surface transmembrane glycoprotein receptor, designated C1qR, as deduced from cDNA (see AA234989) isolated from a 0937 library. C1qR has a predicted pI.wt. of 66,495 and a pI of 5.24. It plays a role in stimulating the classic complement component of the immune system, specifically in stimulating phagocytosis in cells without a concomitant increase in inflammation. Methods for detecting novel ligands for C1qR, including those which function as agonists or antagonists, are provided, as well as methods of determining compositions which effect the formation of an affinity complex between the C1qR and its ligand, and for determining compositions which modulate signal transduction via the C1qR. Transgenic animals can be created to aid in the study of the role of C1qR during growth and metabolism and as a model for disease states in which the normal level of C1qR is effected. The ability to regulate the phagocytic capacity of myeloid cells via the regulation of cell surface expression and function of C1qR will be valuable as a prophylactic treatment of individuals at risk from infection, particularly those with genetic immunodeficiencies, HIV infection, or undergoing cancer chemotherapy or high risk surgery.

Sequence 652 AA:

Query Match 11.8%; Score 307; DB 21; Length 652;
 Best Local Similarity 21.4%; Pred. No. 6.4e-16;
 Matches 139; Conservative 81; Mismatches 206; Indels 224; Gaps 28;

QY 1 MRPAFALCLMQLMPCPGGGEHPADRAGCSAGACYSLSHATMKQAEPACILRGGA 60
 DB 1 matsmgllllllltgpgagtgadteavvcvt-acytahsgklsaaeqnhcngngn 59
 QY 61 LSTRAGAE--LRAVLA--LIRAGPGGGSKDLLFWALERRSHCTLENEPLRGFSW 115
 DB 60 latvkskeagvqvrlaqllrreaaltarmsk---fwlglgkexkclpdlplkgfsw 116
 QY 116 LSSDPGLSDDTLQWVEEQRSCSTARCA--VLQATGVEP--AGWKEMRC-----HL 164
 DB 117 V9---ggedtysnmwkhelrnsckrsvllldlsqpllnrlpkwsegpcspgspgs 173
 QY 165 RANGYLCKYQFVFLCPAPRPGAAASNLRYRAPQLSHALDFSPGTEVSA----- 214
 DB 174 nlegfvckfsfkmgmrplalggpgvtyltptqtsslsleavfasaanvaegedkdet 233
 QY 215 -----LCRQQLP-----ISVTCIADETIGARWD-----KL 238
 DB 234 gshyflckekapdvfdwssgplcvspkygcnfnngchqdcfeggdgsflcgrpgffl 293
 QY 239 SGDVL-----CPGGRY-----LRAGKCAELP 260
 DB 294 lddlvtcasrncpsppcrpgatcwlphgknylrcrpgyqlsdssqldcvdcqcdsp 353
 QY 261 ---NCIDLDLGGFACACATGFE----- 278
 DB 354 caqecvnpmpgfrcecwgyepgpgagacqvdcdcalgrspcagqcnttdgsfnscsee 413

QY 279 ----LCKDGRSC-----VTSGEOP-----TLGCTGYPTRRPPAT 309
 DB 414 gyvlagedgtqcdvdcvpgpplcdslcfnqtgsfhcgclpgwvllapngvscmtmpvs 473
 QY 310 ATSPVQRTMPLRVD--EKLGEPPLVPEQDNSVTSIP-ELPRMGSGSTSTLQMSLQAES 366
 DB 474 lpppsg---pdeedkgekgst--vp---raatastptgpegtphatpstrpslsda 525
 QY 367 KAT-----ITPSGS-----VSKFNSTTSARPOAFDSSA-----VVEI 401
 DB 526 pltsaplklapsgsgvwrrepslhatasgpgpeggdsavatqngtdgklllly 585
 QY 402 EVSTAVVVLITMTVLGLVKLCFHESPSQPRKESMGPPGLESDPEPPA 451
 DB 586 lltvvaillllal-ahgl--lvyrkrakreek-----ekpqnaa 625

Search completed: August 13, 2002, 09:33:29
 Job time: 129 sec

Tue Aug 13 09:56:25 2002

us-09-902-713b-96.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:31:40 ; Search time 16.12 Seconds

(without alignments)
742.465 Million cell updates/sec

Title: US-09-902-713b-96

Perfect score: 2605

Sequence: 1 MRPAFALCLMQLMPCPGC.....LRDRAEGALLAESPLGSSDA 490

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 307 | 11.8 | 652 | 2 | US-08-751-305-2 |
| 2 | 230 | 8.8 | 572 | 6 | 5256770-7 |
| 3 | 227 | 8.7 | 446 | 1 | US-08-307-444A-5 |
| 4 | 227 | 8.7 | 446 | 1 | US-08-587-389-5 |
| 5 | 227 | 8.7 | 456 | 1 | US-08-307-444A-3 |
| 6 | 227 | 8.7 | 456 | 1 | US-08-307-444A-4 |
| 7 | 227 | 8.7 | 456 | 1 | US-08-587-389-3 |
| 8 | 227 | 8.7 | 456 | 1 | US-08-587-389-4 |
| 9 | 227 | 8.7 | 475 | 1 | US-08-307-444A-1 |
| 10 | 227 | 8.7 | 475 | 1 | US-08-307-444A-2 |
| 11 | 227 | 8.7 | 475 | 1 | US-08-587-389-1 |
| 12 | 227 | 8.7 | 475 | 1 | US-08-587-389-2 |
| 13 | 227 | 8.7 | 476 | 1 | US-08-014-723-1 |
| 14 | 227 | 8.7 | 476 | 1 | US-08-014-723-2 |
| 15 | 227 | 8.7 | 476 | 1 | US-08-014-723-18 |
| 16 | 227 | 8.7 | 476 | 1 | US-08-110-011A-1 |
| 17 | 227 | 8.7 | 476 | 1 | US-08-110-011A-2 |
| 18 | 227 | 8.7 | 476 | 1 | US-08-110-011A-18 |
| 19 | 227 | 8.7 | 494 | 1 | US-08-014-723-14 |
| 20 | 227 | 8.7 | 494 | 1 | US-08-014-723-16 |
| 21 | 227 | 8.7 | 494 | 1 | US-08-110-011A-14 |
| 22 | 227 | 8.7 | 494 | 1 | US-08-110-011A-16 |
| 23 | 227 | 8.7 | 497 | 1 | US-08-312-870-3 |
| 24 | 227 | 8.7 | 498 | 2 | US-08-733-564-2 |
| 25 | 227 | 8.7 | 575 | 1 | US-08-261-206A-59 |
| 26 | 227 | 8.7 | 575 | 1 | US-08-312-870-1 |
| 27 | 227 | 8.7 | 575 | 6 | 5466668-6 |

| | | | | | | |
|----|-------|-----|------|---|-------------------|-------------------|
| 28 | 222 | 8.5 | 575 | 1 | US-08-170-290A-54 | Sequence 54, App1 |
| 29 | 175.5 | 6.7 | 215 | 1 | US-08-312-870-5 | Sequence 5, App11 |
| 30 | 133.5 | 5.1 | 443 | 2 | US-08-833-963C-2 | Sequence 2, App11 |
| 31 | 133.5 | 5.1 | 443 | 3 | US-08-980-514-1 | Sequence 1, App11 |
| 32 | 132.5 | 5.1 | 1065 | 2 | US-08-400-159-8 | Sequence 8, App11 |
| 33 | 132.5 | 5.1 | 1212 | 4 | US-09-214-278-3 | Sequence 3, App11 |
| 34 | 132.5 | 5.1 | 1238 | 4 | US-09-214-278-5 | Sequence 5, App11 |
| 35 | 132.5 | 5.1 | 1237 | 4 | US-08-611-729A-8 | Sequence 8, App11 |
| 36 | 130.5 | 5.0 | 985 | 5 | PCT-US96-03916-66 | Sequence 6, App11 |
| 37 | 130.5 | 5.0 | 985 | 5 | PCT-US96-03916-66 | Sequence 6, App11 |
| 38 | 126.5 | 4.9 | 1055 | 4 | US-09-214-278-2 | Sequence 2, App11 |
| 39 | 124 | 4.8 | 933 | 2 | US-08-313-200-1 | Sequence 1, App11 |
| 40 | 124 | 4.8 | 933 | 5 | PCT-US93-03837-1 | Sequence 1, App11 |
| 41 | 123 | 4.7 | 1148 | 4 | US-08-882-046-4 | Sequence 4, App11 |
| 42 | 122.5 | 4.7 | 1479 | 3 | US-08-840-062-2 | Sequence 2, App11 |
| 43 | 121.5 | 4.7 | 1455 | 3 | US-08-840-062-5 | Sequence 5, App11 |
| 44 | 118.5 | 4.5 | 1404 | 2 | US-08-400-159-2 | Sequence 2, App11 |
| 45 | 118.5 | 4.5 | 1404 | 3 | US-08-611-729A-2 | Sequence 2, App11 |

ALIGNMENTS

RESULT: 1
US-08-751-305-2
; Sequence 2, Application US/08751305
; Patent No. 5965439
; GENERAL INFORMATION:
; APPLICANT: Tennet et al., Andrea J.
; TITLE OF INVENTION: HOST DEFENSE ENHANCEMENT
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,305
; FILING DATE: 18-NOV-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07306/012001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-751-305-2

Query Match 11.8%; Score 307; DB 2; Length 652;
Best Local Similarity 21.4%; Pred. No. 2.4e-17;
Matches 139; Conservative 81; Mismatches 206; Indels 224; Gaps 28;

QY 1 MRPAFALCLMQLMPCPGCEHPTRADACCSASGACTSLHATMKRQAAEEACILRGCA 60
DB 1 MATSMGLILLILLITLTPGAGTGADTEAVVCGT-ACVTAHSKLSAEEAQNONGN 59
QY 61 LSTVRAAE---LRAVLA--LIRAGPDPGGGSKDLFLFWALERRRRSICLTLENPLRGSF 115

Db 60 LATVSKKEAOHYORVLAOLLRREALTARMK--FWIGLOREKKCLDPSLPLGFSN 116
 QY 116 LSSPGLESDTLQWVEEPQSTARGA--VQATGVEP--ACWKEMRC-----HL 164
 Db 117 VG--GGEDLP--GNMHEKLNKSTIKRCVSLDLISQPLIPNLPKPMSEPCGSPGSPGS 173
 QY 165 RANGVLCYKOFELCPAPRGASNLSTYRPFQJHSAALDFSPGTEVSA-----214
 Db 174 NIEGFVKESFKMCRPLALGCGVLYTTPFQTTSSILEVFPASAAVACGECDKET 233
 QY 215 -----LCRGDLP-----ISVTCIDELGARD-----KL 238
 Db 234 QSHYFLCKEKAPVPYFMSSGGLCVSPKYGKCNFNNGGCHODCFEGDGSFLGCRPGFRL 293
 QY 239 SGDVL-----CPCPRY-----LAGCAELP 260
 Db 294 LDDLVLCAARNPCSSPFCGATCVLGHGKNTYTCGCPGYDLDSDLDVYDCQDSP 353
 QY 261 ----NCDDLGFAECATGFE-----278
 Db 354 CAQECVNTGGRFCRCWGWIEGCGEGACQDVCALCRSPCAQCTNTDGFHCSEE 413
 QY 279 ----LKDGRSC-----VTSGEQGP-----TLGGVPTTRPPAT 309
 Db 414 GYVLAGEOSTQGDYDECVPGRPLCDSLCFNTQSFHGGCLPGVYLAIPNGVSTMGFVS 473
 QY 310 ATSPVPQRTPIKVD--EKLGTEPLVPQDQNSVTIP--EIPRMGSQSTMTLQMSLOAES 366
 Db 474 LGPPGSP--PDEEDKGEKEGT--VP--RAATASPRGEGPKAIPPTTSRSLSDA 525
 QY 367 KAT-----TPSGS-----VISKSTISSATPOAFDSSA-----VET 401
 Db 526 PITSAPLKLAPSSSSGWRPDSIHATAGPOEPAGDSSVAATQNDGTDGKLLIFY 585
 QY 402 FVSTAVVVLITIMVGLVLCHESSPSQPKRESMGPGLSEDPEDPA 451
 Db 586 ILGTVAALLLAL--ALGL--LVYRKRRAKKEKK-----EKKPQMA 625

RESULT 2
 5256770-7
 Patent No. 5256770
 APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT, DAVID R.
 TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
 NUMBER OF SEQUENCES: 48
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/506,325
 FILING DATE: 09-APR-1990
 ID NO: 7
 LENGTH: 572
 5256770-7

Query Match 8.8%; Score 230; DB 6; Length 572;
 Best Local Similarity 26.3%; Pred. No. 4,9e-11;
 Matches 78; Conservative 38; Mismatches 119; Indels 62; Gaps 15;
 QY 16 PGGGGEHPYADRAAGCSAGACSLHATMKROAEACILRGALSTVRAGAEIARVLA 75
 Db 23 PGGGSGCVVHD-----CFALYGPAPFLNMAQICDGLRLHMTVRSSVAADVISL 73
 QY 76 LIRAGPGGSGKDLFWALE-----RRSHCTLENEPLRGFSWSSDPGLESPT 127
 Db 74 LLN---GDGVGRRRL--WIGLQLPFGCGDPRRG-----PLRGFWYTGDNNTSYS-- 119
 QY 128 LQWVEEPQRS-----CTARCAVILQATGVEPAGWKEMRCHLRANGYLCKYQFVLCP 180
 Db 120 -RWARLDLNGAPLGGPLCAVVSAA--EATVPSEPI--WEOQCEYKADGLFCEHFPAICR 175
 QY 181 -APRGAAS--NLSTYRAPFOLHSAALDFSPGTEVSALCGQPLISTYTCIADETGAW 235
 Db 176 PLAVEGAAALAAVSTYTGPFARAGADFOALPVGSSAAV--APGLQIMCTAGNVQGHW 232

QY 236 DKLSDVLCPGGRY--LRAGKCAELPNCDDLGCFACGATGFEKDGSRCTVS 289
 Db 223 AR-----EAPGAMDCSVENGCEHACNAIP--GAPRCQCPAGAAIADGRSCTAS 280

RESULT 3
 US-08-307-444A-5
 Sequence 5, Application US/08307444A
 Patent No. 5516559
 GENERAL INFORMATION:
 APPLICANT: NII, ATSUSHI
 APPLICANT: MORISHITA, HIDEAKI
 APPLICANT: UEMURA, AKIO
 APPLICANT: MOCHIDA, EI
 TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OLIFF & BERRIDGE
 STREET: P.O. BOX 19928
 CITY: ALEXANDRIA
 STATE: VA
 COUNTRY: USA
 ZIP: 22320
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/307,444A
 FILING DATE: 19-SEP-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/835,436
 FILING DATE: 26-FEB-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: OLIFF, JAMES A.
 REGISTRATION NUMBER: 27,075
 REFERENCE/DOCKET NUMBER: JAO 27706.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6400
 TELEFAX: (703) 836-2787
 TELEX: 90-1799 PTO ALEX
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 446 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-307-444A-5

Query Match 8.7%; Score 227; DB 1; Length 446;
 Best Local Similarity 25.7%; Pred. No. 6.2e-11;
 Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;
 QY 16 PGGGGEHPYADRAAGCSAGACSLHATMKROAEACILRGALSTVRAGAEIARVLA 75
 Db 5 PGGGSGCVVHD-----CFALYGPAPFLNMAQICDGLRLHMTVRSSVAADVISL 55
 QY 76 LIRAGPGGSGKDLFWALE-----RRSHCTLENEPLRGFSWSSDPGLESPT 127
 Db 56 LLN---GDGVGRRRL--WIGLQLPFGCGDPRRG-----PLRGFWYTGDNNTSYS-- 102
 QY 128 LQWVEEPQRS-----CTARCAVILQATGVEPAGWKEMRCHLRANGYLCKYQFVLCP 180
 Db 103 -RWARLDLNGAPLGGPLCAVVSAA--EATVPSEPI--WEOQCEYKADGLFCEHFPAICR 158
 QY 181 -APRGAAS--NLSTYRAPFOLHSAALDFSPGTEVSALCGQPLISTYTCIAD-----229

Db 159 PLAVEGAAAAAIVITYGTFPARGADFOALPVGSSAAV---APLGIQLMCTAPPGAVOG 215
QY 230 -----EIGARD-----KLSGDLCP-----GRYLRAG---KCAEL--- 259
Db 216 HWAREAPGAMDCSVENGCGEHCACNAIPGAPRCQCPAGALQADGRSCTASATOSCNDLCE 275
QY 260 ----PNCIDLDGGFACGATGFEIGKDGRC 286
Db 276 HFCVFN-PDPGSGYSCMCEGTGYRLADQHRC 305

RESULT 4

US-08-587-389-5
; Sequence 5, Application US/08587389
; Patent No. 5695964

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO

APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT

TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/587,389

FILING DATE: 17-JAN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/307,444

FILING DATE: 19-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: OLIF, JAMES A.

REGISTRATION NUMBER: 27,075

REFERENCE/DOCKET NUMBER: JAO 27706

TELEPHONE: (703) 836-6400

TELEFAX: (703) 836-2787

TELEX: 90-1799 PTO ALEX

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 446 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-587-389-5

Query Match 8.7%; Score 227; DB 1; Length 446;

Best Local Similarity 25.7%; Pred. No. 6,2e-11;

Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PGFGGEHPTADRGASGACYSLSHATWKQAAEACTLRGALSTVRAGELRAVLA 75

Db 5 PDPGSGYSCMCEGTGYRLADQHRC 305

QY 76 LIRAGGPGGSGDLFWALE-----RRSHCTLENEPLRGSSWLSDDGGLSDT 127

Db 56 LFN---GDGSGRRRL-WIGLQLEPGCGDKRLG-----PLRGFWWTGDNNTSYS-- 102
QY 128 LOWVEEQR-----CTARCAVLQATGVEPAGKEMRCHLRANGYLCKTQFVLC 180
Db 103 -RRARLDINAPLCGLCAVNSAA--EATVPSERI-WEEQCEYKADGFLCEHPATCR 158
QY 181 -APRGAAS--NLSYRPFQHSALDPSPPCTEVSALCRGOLPISTYCIAD----- 229
Db 159 PLAVEGAAAAAIVITYGTFPARGADFOALPVGSSAAV---APLGIQLMCTAPPGAVOG 215
QY 230 -----EIGARD-----KLSGDLCP-----GRYLRAG---KCAEL--- 259
Db 216 HWAREAPGAMDCSVENGCGEHCACNAIPGAPRCQCPAGALQADGRSCTASATOSCNDLCE 275
QY 260 ----PNCIDLDGGFACGATGFEIGKDGRC 286
Db 276 HFCVFN-PDPGSGYSCMCEGTGYRLADQHRC 305

RESULT 5

US-08-307-444A-3
; Sequence 3, Application US/08307444A
; Patent No. 5516659

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO

APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/307,444A

FILING DATE: 19-SEP-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/835,436

FILING DATE: 26-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: OLIF, JAMES A.

REGISTRATION NUMBER: 27,075

REFERENCE/DOCKET NUMBER: JAO 27706

TELEPHONE: (703) 836-6400

TELEFAX: (703) 836-2787

TELEX: 90-1799 PTO ALEX

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-307-444A-3

Query Match 8.7%; Score 227; DB 1; Length 456;

Best Local Similarity 25.7%; Pred. No. 6,4e-11;

Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PGGGGEHPADRGCSAGSACYSIHATMKROAEACILRGALSTVRAGELRAVLA 75
DB 5 POPGSGQVCEHD-----CFALYGPATFTLNASQICDGLGHMTVRSSVAADVLSL 55
QY 76 LLRAGPBGSGSKDLFWALF-----RRSHCTLENEPLRGFSWLSDPGGLSDPT 127
DB 56 LLN---GDGSGVGRRL-WIGLOLPBGCGDPKRLG-----PLRGQVWTGNNNTSYS-- 102
QY 128 LOWEERFORS-----CTARRCAVLOATGVGPACMKRMCHLRANGTLCKYOEVLCP 180
DB 103 -RWARLDLNCAPLCPLCAVASNA--EATVPSEPT-WBEOQCEVADGFLCEHFPATCR 158
QY 181 -APRGAS--NLSTRAPFOLHSALDPSPTGEVSALCRGOLPISTVCTAD----- 229
DB 159 PLAVEGMAAAVSTYGTPEFARGADFOALPVGSSAAV--APLGLOMCTAPPGAVOG 215
QY 230 ---EIGARD-----KLSDVLCPCP-----GRLRAG--KCAEL-- 259
DB 216 HMARERAPGAMDCSVENGGEHACNAIPGAPRCQCPAGALADGRSCTASATQSCNDICE 275
DB 260 ---PNCIDLDGFACECATGFEIGKDGRC 286
DB 276 HRCVBN-PDQPSYSQCMCTGYRLAADQHC 305

RESULT 6

US-08-307-444A-4

Sequence 4, Application US/08307444A

Patent No. 551659

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI

APPLICANT: MORISHITA, HIDEAKI

APPLICANT: UEMURA, AKIO

APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/307,444A

FILING DATE: 19-SEP-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/835,436

FILING DATE: 26-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: OLIFF, JAMES A.

REGISTRATION NUMBER: 27, 075

REFERENCE/DOCKET NUMBER: JAO 27706

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6400

TELEFAX: (703) 836-2787

TELEX: 90-1799 PTO ALEX

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

STRANDEDNESS: single

MOLECULE TYPE: protein

HYPOTHEICAL: NO

ANTI-SENSE: NO

US-08-307-444A-4

Query Match 8.7%; Score 227; DB 1; Length 456;
Best Local Similarity 25.7%; Pred. 6.4e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PGGGGEHPADRGCSAGSACYSIHATMKROAEACILRGALSTVRAGELRAVLA 75
DB 5 POPGSGQVCEHD-----CFALYGPATFTLNASQICDGLGHMTVRSSVAADVLSL 55
QY 76 LLRAGPBGSGSKDLFWALF-----RRSHCTLENEPLRGFSWLSDPGGLSDPT 127
DB 56 LLN---GDGSGVGRRL-WIGLOLPBGCGDPKRLG-----PLRGQVWTGNNNTSYS-- 102
QY 128 LOWEERFORS-----CTARRCAVLOATGVGPACMKRMCHLRANGTLCKYOEVLCP 180
DB 103 -RWARLDLNCAPLCPLCAVASNA--EATVPSEPT-WBEOQCEVADGFLCEHFPATCR 158
QY 181 -APRGAS--NLSTRAPFOLHSALDPSPTGEVSALCRGOLPISTVCTAD----- 229
DB 159 PLAVEGMAAAVSTYGTPEFARGADFOALPVGSSAAV--APLGLOMCTAPPGAVOG 215
QY 230 ---EIGARD-----KLSDVLCPCP-----GRLRAG--KCAEL-- 259
DB 216 HMARERAPGAMDCSVENGGEHACNAIPGAPRCQCPAGALADGRSCTASATQSCNDICE 275
DB 260 ---PNCIDLDGFACECATGFEIGKDGRC 286
DB 276 HRCVBN-PDQPSYSQCMCTGYRLAADQHC 305

RESULT 7

US-08-587-389-3

Sequence 3, Application US/08587389

Patent No. 5695964

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI

APPLICANT: MORISHITA, HIDEAKI

APPLICANT: UEMURA, AKIO

APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/587,389

FILING DATE: 17-JAN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/307,444

FILING DATE: 19-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: OLIFF, JAMES A.

REGISTRATION NUMBER: 27, 075

REFERENCE/DOCKET NUMBER: JAO 27706

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6400

TELEFAX: (703) 836-2787

TELEX: 90-1799 PTO ALEX

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

US-08-587-389-3

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-587-389-3

Query Match 8.7%; Score 227; DB 1; Length 456;
Best Local Similarity 25.7%; Pred. No. 6.4e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PPGGGEHPTADRCASGACYSIHHATMKROAAEACILRGALSTVRAGAEI RAVLA 75
DB 5 PPGGSGQCEVHD-----CFALYGPATFLNMQICDGLRHLMTVRSSVADVLSL 55
QY 76 LLRAGPGGGSKDILFWVALE-----RRSHCTLENEPLRGFSMLSPDGLSEPT 127
DB 56 LLN---GGGCGVRRRL-WIGLQLPFGCGDPKRLG-----PLRGFWVTGDNNTSYS-- 102
QY 128 LQWVEEPPORS-----CTARCAVLQATGVEPAGMKEMRCHLRANGYLCKYQFEVLCP 180
DB 103 -RMARLDLNGAPLCGLCVAVSA--EATVSEPT-WEEQCEVAKADGFLCEHFPATCH 158
QY 181 -APRPGAS--NLSTYAPFQLHSAALDFSPGTEVSALCRGOLPISTCIAD----- 229
DB 159 PLAVEPAGAAVAASITYGTFPARGADFOALPVGSSAAV--APLGLQIMCTAPPGAVOG 215
QY 230 -ELGARMD-----KLSDVLCPCP-----GRLVRG---KKAEL-- 259
DB 216 HWAREPAGMDCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCE 275
QY 260 ---PNCDDLDGFACCECATGFEIGKDGRC 286
DB 276 HFCVPR-PDQPGSYSCMCTGYRLAADQHR 305

RESULT 8
US-08-587-389-4
Sequence 4, Application US/08587389
Patent No. 5695964

GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCHIDA, EI
TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,389
FILING DATE: 17-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/307,444
FILING DATE: 19-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: OLIFF, JAMES A.
REGISTRATION NUMBER: 27,075
REFERENCE/DOCKET NUMBER: JAO 27706

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
TELEX: 90-1799 PTO ALEX
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-587-389-4

Query Match 8.7%; Score 227; DB 1; Length 456;
Best Local Similarity 25.7%; Pred. No. 6.4e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PPGGGEHPTADRCASGACYSIHHATMKROAAEACILRGALSTVRAGAEI RAVLA 75
DB 5 PPGGSGQCEVHD-----CFALYGPATFLNMQICDGLRHLMTVRSSVADVLSL 55
QY 76 LLRAGPGGGSKDILFWVALE-----RRSHCTLENEPLRGFSMLSPDGLSEPT 127
DB 56 LLN---GGGCGVRRRL-WIGLQLPFGCGDPKRLG-----PLRGFWVTGDNNTSYS-- 102
QY 128 LQWVEEPPORS-----CTARCAVLQATGVEPAGMKEMRCHLRANGYLCKYQFEVLCP 180
DB 103 -RMARLDLNGAPLCGLCVAVSA--EATVSEPT-WEEQCEVAKADGFLCEHFPATCH 158
QY 181 -APRPGAS--NLSTYAPFQLHSAALDFSPGTEVSALCRGOLPISTCIAD----- 229
DB 159 PLAVEPAGAAVAASITYGTFPARGADFOALPVGSSAAV--APLGLQIMCTAPPGAVOG 215
QY 230 -ELGARMD-----KLSDVLCPCP-----GRLVRG---KKAEL-- 259
DB 216 HWAREPAGMDCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCE 275
QY 260 ---PNCDDLDGFACCECATGFEIGKDGRC 286
DB 276 HFCVPR-PDQPGSYSCMCTGYRLAADQHR 305

RESULT 9
US-08-307-444A-1
Sequence 1, Application US/08307444A
Patent No. 5516659

GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCHIDA, EI
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,444A
FILING DATE: 19-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/835,436
FILING DATE: 26-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: OLIFF, JAMES A.
REGISTRATION NUMBER: 27,075
REFERENCE/DOCKET NUMBER: JAO 27706
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
TELEX: 90-1799 PRO ALEX
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
307-444A-1

Query Match 8.7%; Score 227; DB 1; Length 475;
Best Local Similarity 25.7%; Pred. No. 6,7e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PGGGGEHPTADRCASGACYSLHATMKRQAEACILRGALSTVRAGELRAVLA 75
DB 23 PPGGSGQVEHD-----CFALYPGPATFLNLSQICGLHMTVRSVAADYISL 73
QY 76 LRRGPGGGSKDLFWALE-----RRSHCTLENEPLRGFSWLSDDPGLESST 127
DB 74 LLN---GDGVGRRRL-WIGLQLEPPGCGDKPRLG-----PLRGQWVGDNNTSYS-- 120
QY 128 LQWVEEPORS-----CTARCAVLQATGVEPAGKEMKCHLRANGYLCKYQFEVLC 180
DB 121 -RWARLDNGAPLGPCLVAVSA--EATVPSPI-WEEQCEVKADGFLCEHFPAICR 176
QY 181 -APRGAAS--NLSTAPRQLHSAALDPSPPCTEVSALCRGLPISTVICAD----- 229
DB 177 PLAVEPAAAAAASITTYGTFFAARGADFOALPVGSSAAV---APLGLQIMCTAPPGAVOG 233
QY 230 ---EIGARDW-----KLSGDVLCPCP-----GRYLKAG--KCAEL-- 259
DB 234 HWAREAPGAMDCSVENGCGEHCACNAIPGAPRCQCPAGALADGRSCTASATQSCNDLCE 293
QY 260 ---PNCDDLGSPACACATGFELEGKDGKRS 286
294 HFCVPM-PDOPGYSYSCMCEGYRLAADQHR 323

RESULT 10
US-08-307-444A-2
Sequence 2, Application US/08307444A
Patent No. 551659

GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCHIDA, EI
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,444A
FILING DATE: 19-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/835,436
FILING DATE: 26-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: OLIFF, JAMES A.
REGISTRATION NUMBER: 27,075
REFERENCE/DOCKET NUMBER: JAO 27706
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
TELEX: 90-1799 PRO ALEX
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-307-444A-2

Query Match 8.7%; Score 227; DB 1; Length 475;
Best Local Similarity 25.7%; Pred. No. 6,7e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

QY 16 PGGGGEHPTADRCASGACYSLHATMKRQAEACILRGALSTVRAGELRAVLA 75
DB 23 PPGGSGQVEHD-----CFALYPGPATFLNLSQICGLHMTVRSVAADYISL 73
QY 76 LRRGPGGGSKDLFWALE-----RRSHCTLENEPLRGFSWLSDDPGLESST 127
DB 74 LLN---GDGVGRRRL-WIGLQLEPPGCGDKPRLG-----PLRGQWVGDNNTSYS-- 120
QY 128 LQWVEEPORS-----CTARCAVLQATGVEPAGKEMKCHLRANGYLCKYQFEVLC 180
DB 121 -RWARLDNGAPLGPCLVAVSA--EATVPSPI-WEEQCEVKADGFLCEHFPAICR 176
QY 181 -APRGAAS--NLSTAPRQLHSAALDPSPPCTEVSALCRGLPISTVICAD----- 229
DB 177 PLAVEPAAAAAASITTYGTFFAARGADFOALPVGSSAAV---APLGLQIMCTAPPGAVOG 233
QY 230 ---EIGARDW-----KLSGDVLCPCP-----GRYLKAG--KCAEL-- 259
DB 234 HWAREAPGAMDCSVENGCGEHCACNAIPGAPRCQCPAGALADGRSCTASATQSCNDLCE 293
QY 260 ---PNCDDLGSPACACATGFELEGKDGKRS 286
294 HFCVPM-PDOPGYSYSCMCEGYRLAADQHR 323

RESULT 11
US-08-587-389-1
Sequence 1, Application US/08587389
Patent No. 569564

GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCHIDA, EI
TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: ALEXANDRIA

```

1 STATE: VA
2 COUNTRY: USA
3 ZIP: 22320
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 OPERATING SYSTEM: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patent In Release #1.0, Version #1.25
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/08/587,389
11 FILING DATE: 17-JAN-1996
12 CLASSIFICATION: 435
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 08/307,444
15 FILING DATE: 19-SEP-1994
16 ATTORNEY/AGENT INFORMATION:
17 NAME: OLIFE, JAMES A.
18 REGISTRATION NUMBER: 27,075
19 REFERENCE/DOCKET NUMBER: JO 27706
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (703) 836-6400
22 TELEFAX: (703) 836-2787
23 TELE: 90-1799 PRO ALEX
24 INFORMATION FOR SEQ ID NO: 1:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH: 475 amino acids
27 TYPE: amino acid
28 STRANDEDNESS: single
29 TOPOLOGY: linear
30 MOLECULE TYPE: protein
31 HYPOTHEICAL: NO
32 ANTI-SENSE: NO
33 US-08-587-389-1
34
35 Query Match 8.7% Score 227; DB 1; Length 475;
36 Best Local Similarity 25.7%; Pred. No. 6.7e-11;
37 Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18
38
39 Oy 16 PGGGGEHTADRACCSAGACYSYSHATMKROAEACILRGALSTVRAGAEIRAVIA 75
40 Db 23 PPGGSGQCEHD-----CFALYPGATPLNNSQIDGLRGLMTVRSYAAAVISL 73
41 Oy 76 LIRAGPREGGSKDILFWALE-----RRRSHCTLENEPLRGFSWLSDPGGLESPT 127
42 Db 74 ILN---GGGVGRRRL-WIGLQLPPGCCGPKRLG-----PLRGQVWTGDNNTSYS-- 120
43 Oy 128 LOWVEPQRS-----CTARCAVLQATGVGPACWKMKRCHLRANGILCKYQFVLCIP 180
44 Db 121 -RWARLDLNGAPLCPILCAVASAA--EATVPSEPT-WEQDCEVAKDGLCEHFHPATCR 176
45 Oy 181 -APRGAAS--NISYRAPFOLHSAALDFSPGTEVSALRGQLPISVTCIAD----- 229
46 Db 177 PLAVEPGAAAAAVSTTYGPFARAGADPQALFVGSSAAV--APLGLQLMCTARPPAVOG 233
47 Oy 230 ---EIGARWD-----KLSDGVLCPCP-----GRYLARG--KCAEL-- 259
48 Db 234 HWAREAPGAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRCTASATQSCNDLCE 293
49 Oy 260 ---PNCDDLDGGFACECATGFEIGKDGASC 286
50 Db 294 HFCVFN-PDQPGSYSCMCETGYRLAADQHRC 323
51
52 RESULT 12
53 US-08-587-389-2
54 Sequence 2, Application US/08587389
55 Patent No. 5695364
56 GENERAL INFORMATION:
57 APPLICANT: NII, ATSUSHI
58 APPLICANT: MORISHITA, HIDEAKI
59 APPLICANT: UEMURA, AKIO
60 APPLICANT: MOCHIDA, ET

```

Sequence 1, Application US/08014723
Patent No. 5273962

GENERAL INFORMATION:

APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Okuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014,723
FILING DATE: 19930208

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5273962man F.
REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 80-071-0 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)413-3000

TELEFAX: (703)413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-014-723-1

Query Match 8.7% Score 227; DB 1; Length 476;
Best Local Similarity 25.7%; Pred. No. 6.8e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

16 PGGGSEHPTADRGASGACYSLSHATMRQAEEACILRGALSTVRGAELRAVLA 75
5 PGGGSCQVEHD-----CFALYPPATFLNASQIDGLRGLMTVRSSADVLSL 55
76 ILRAGPGGGSKDLLFWALE-----RRSHCTLENEPLRGFSWLSDDPGLESPT 127
56 LLN---GGGAGVGRRL-WIGLDLPCCGDPKRLG-----PLRGFWMTGDNNTSYS-- 102
128 LOWEPEPORS-----CTARCAVLATGATGVEPAGMKEMKCHIRANGYLCKYQEVLC 180
103 -RVARLDLNGAPLCCPLCAVSA--EATVSEPT-WEEOQCEVKADGFLCEHFPACTR 158
181 -APRGAAS--NLSTAPFOLHSAALDFSPGTEVSALCRGQLPISTVCID----- 229
159 PLAVEPAAAAAASITVGTFFAARGADFOALPVSSAAV--APLGQLMCTAPPGAVOG 215
230 -----ETGARMD-----KLSGDLVLCPP-----GRYLARG--KCAEL-- 259
216 HWAAREPAGMDCSVENGCEHACNAIPGAPRCOCPPAGALADGRSCTASATQSCNDICE 275
260 -----PNCIDLDGFGACACATGFEIGKDGRC 286
276 HFCVNP-PDQPGSYSCMCEGTGRLAADQHR 305

RESULT 14
US-08-014-723-2

Sequence 2, Application US/08014723
Patent No. 5273962

GENERAL INFORMATION:

APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Okuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014,723
FILING DATE: 19930208

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5273962man F.
REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 80-071-0 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)413-3000

TELEFAX: (703)413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-014-723-2

Query Match 8.7% Score 227; DB 1; Length 476;
Best Local Similarity 25.7%; Pred. No. 6.8e-11;
Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

16 PGGGSEHPTADRGASGACYSLSHATMRQAEEACILRGALSTVRGAELRAVLA 75
5 PGGGSCQVEHD-----CFALYPPATFLNASQIDGLRGLMTVRSSADVLSL 55
76 ILRAGPGGGSKDLLFWALE-----RRSHCTLENEPLRGFSWLSDDPGLESPT 127
56 LLN---GGGAGVGRRL-WIGLDLPCCGDPKRLG-----PLRGFWMTGDNNTSYS-- 102
128 LOWEPEPORS-----CTARCAVLATGATGVEPAGMKEMKCHIRANGYLCKYQEVLC 180
103 -RVARLDLNGAPLCCPLCAVSA--EATVSEPT-WEEOQCEVKADGFLCEHFPACTR 158
181 -APRGAAS--NLSTAPFOLHSAALDFSPGTEVSALCRGQLPISTVCID----- 229
159 PLAVEPAAAAAASITVGTFFAARGADFOALPVSSAAV--APLGQLMCTAPPGAVOG 215
230 -----ETGARMD-----KLSGDLVLCPP-----GRYLARG--KCAEL-- 259
216 HWAAREPAGMDCSVENGCEHACNAIPGAPRCOCPPAGALADGRSCTASATQSCNDICE 275
260 -----PNCIDLDGFGACACATGFEIGKDGRC 286
276 HFCVNP-PDQPGSYSCMCEGTGRLAADQHR 305

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:32:00 ; Search time 20.93 Seconds

(without alignments)
2249.583 Million cell updates/sec

Title: US-09-902-713B-96

Perfect score: 2605

Sequence: 1 MRPAFALCLMQALWPGPG.....LRDRAEGALLAESPLGSSDA 490

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 227 | 8.7 | 575 | 1 | THHUB |
| 2 | 192 | 7.4 | 577 | 2 | thrombomodulin pre |
| 3 | 132 | 5.1 | 3507 | 2 | thrombomodulin pre |
| 4 | 129.5 | 5.0 | 1106 | 2 | hypothetical prote |
| 5 | 124.5 | 4.8 | 2477 | 2 | hypothetical prote |
| 6 | 122.5 | 4.7 | 1479 | 2 | fibronectin precu |
| 7 | 122 | 4.7 | 933 | 2 | mannose receptor, |
| 8 | 122 | 4.7 | 2471 | 2 | iodide peroxidase |
| 9 | 121.5 | 4.7 | 462 | 2 | cell-fate determin |
| 10 | 121.5 | 4.7 | 1455 | 1 | hypothetical prote |
| 11 | 119 | 4.6 | 646 | 2 | mannose receptor p |
| 12 | 119 | 4.6 | 2907 | 2 | fibronectin precu |
| 13 | 118.5 | 4.5 | 1408 | 2 | fibronectin-2 precu |
| 14 | 118.5 | 4.5 | 3020 | 2 | gene serrate prote |
| 15 | 117.5 | 4.5 | 2318 | 2 | mucin 2 precursor, |
| 16 | 117 | 4.5 | 2871 | 2 | notch 3 protein - |
| 17 | 117 | 4.5 | 2871 | 2 | fibronectin-1 precu |
| 18 | 117 | 4.5 | 3002 | 2 | fibronectin-1 precu |
| 19 | 117 | 4.5 | 3570 | 2 | fibronectin-1 precu |
| 20 | 116 | 4.5 | 2918 | 2 | mucin MUC5B, trach |
| 21 | 115.5 | 4.4 | 1220 | 2 | fibronectin-2 precu |
| 22 | 115 | 4.4 | 1276 | 2 | jagged protein pre |
| 23 | 115 | 4.4 | 4753 | 2 | SRBP cleavage act |
| 24 | 114.5 | 4.4 | 1456 | 1 | LDL-receptor relat |
| 25 | 114.5 | 4.4 | 1964 | 2 | mannose receptor p |
| 26 | 113.5 | 4.4 | 676 | 2 | notch4 - mouse |
| 27 | 113.5 | 4.4 | 2254 | 2 | plasma protein S p |
| 28 | 113.5 | 4.4 | 2910 | 2 | low voltage-activa |
| 29 | 113 | 4.3 | 1612 | 2 | otogelin - mouse |
| | | | | | ducl protein - mo |

| | | | | | | |
|----|-------|-----|------|---|--------|---------------------|
| 30 | 113 | 4.3 | 2139 | 2 | A35672 | crumbs protein - f |
| 31 | 112.5 | 4.3 | 909 | 1 | ORXLL1 | LDL receptor 1 pre |
| 32 | 112.5 | 4.3 | 4660 | 2 | T42737 | gp330 protein prec |
| 33 | 112 | 4.3 | 1221 | 2 | A49457 | fibulin-2 precursor |
| 34 | 112 | 4.3 | 1952 | 2 | T48814 | hypothetical prote |
| 35 | 111 | 4.3 | 2437 | 2 | S42612 | transmembrane prot |
| 36 | 110.5 | 4.2 | 1376 | 2 | G00043 | osteonidogen - hum |
| 37 | 110.5 | 4.2 | 1722 | 2 | E89753 | protein Flic7.4 11 |
| 38 | 110 | 4.2 | 675 | 1 | KXMS | plasma protein S p |
| 39 | 110 | 4.2 | 870 | 2 | T31795 | hypothetical prote |
| 40 | 110 | 4.2 | 1032 | 2 | T34433 | hypothetical prote |
| 41 | 110 | 4.2 | 1574 | 2 | T13954 | MEG6 protein - ra |
| 42 | 110 | 4.2 | 2232 | 2 | T34434 | hypothetical prote |
| 43 | 109.5 | 4.2 | 642 | 2 | S53434 | plasma protein S p |
| 44 | 109.5 | 4.2 | 926 | 1 | OPPGIT | iodide peroxidase |
| 45 | 109 | 4.2 | 675 | 1 | KXRTS | plasma protein S p |

ALIGNMENTS

RESULT 1

THHUB
thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 15-Sep-2000

C:Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruy

J. Biochem. 103, 281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed ac

A:Reference number: A41442; MOID:88227901

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SH1>

A:Cross-references: DDBJ:000210; NID:g220126; PIDN:BAA00149.1; PID:g220127

R:Packman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the

A:Reference number: A28307; MOID:87317665

A:Accession: A28307

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02873; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, T.; Zushi, M.; Kaw

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on end

A:Reference number: A29680; MOID:88004395

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SU2>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: Lung endothelium

A:Note: Part of this sequence, including the amino end of the mature protein, were de

R:Men, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of

A:Reference number: A27073; MOID:88024950

A:Accession: A27073

A:Molecule type: DNA

A:Residues: 1-472, 'A', 474-575 <MEN>

A:Cross-references: GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:g339657

A:Experimental source: Placenta

A:Note: Parts of this sequence were determined by protein sequencing

J:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.

R:Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MOID:93293792

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.

Biochem. J. 295, 131-140, 1993
 A>Title: Identification of the predominant glycosaminoglycan attachment site in soluble serine.
 A:Reference number: S38954; MUID:94029900
 A:Accession: S38954
 A:Molecule type: protein
 A:Residues: 475-491, 'X', 493-494 <GER>
 A>Note: The residue designated 'X' was determined to be a Ser with covalently bound chor
 R.Meinhart, D.P.; Komives, E.A.
 Submitted to the Brookhaven Protein Data Bank, September 1995
 A:Reference number: A67369; PDB:1ZAG
 A:Contents: annotation: conformation and disulfide bond assignments by (1)H-NMR, residue
 R.Tulinsky, A.; Mathews, I.I.
 Submitted to the Brookhaven Protein Data Bank, August 1994
 A:Reference number: A52804; PDB:1HLT
 A:Contents: annotation: X-ray crystallography, 3.0 angstroms, residues 426-442
 R.Hirsh, R.; Komives, E.A.; Ni, F.
 Submitted to the Brookhaven Protein Data Bank, November 1995
 A:Reference number: A65583; PDB:1FGD
 A:Contents: annotation: conformation by (1)H-NMR, residues 427-444
 R.Hirsh, R.; Komives, E.A.; Ni, F.
 Submitted to the Brookhaven Protein Data Bank, November 1995
 A:Reference number: A58595; MUID:96276211
 A:Contents: annotation: conformation by (1)H-NMR
 A:Gene: GDB:THRD
 A:Cross-references: GDB:119613; OMIM:188040
 A:Map position: 20p11.2-20p11.2
 A:Introns: #status absent
 A:Complex: homodimer, urinary form
 C:Function:
 A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activat
 A:Pathway: blood coagulation/moderation
 A>Note: the membrane-bound form is located on the endothelium luminal surface of arterie
 A>Note: Chromomodulin complexed with the membrane-bound form is subject to endocytosis
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coag
 e protein
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
 F:19-513/Domain: extracellular #status predicted <EXT>
 F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>
 F:24-167/Domain: C-type lectin homology <LCH>
 F:177-199/Region: PEST sequence
 F:201-233/Region: PEST sequence
 F:245-280/Domain: EGF homology <EG1>
 F:288-323/Domain: EGF homology <EG2>
 F:362-362/Domain: EGF homology <EG3>
 F:404-439/Domain: EGF homology <EG4>
 F:445-480/Domain: EGF homology <EG5>
 F:485-513/Region: PEST sequence
 F:517-539/Domain: transmembrane #status predicted <TMN>
 F:540-575/Domain: intracellular #status predicted <INT>
 F:47,115,116,382,409/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:174,225,411,504/Binding site: carboxylate (Thr) (covalent) #status predicted
 F:245-256,252-265,292-308,310-323,329-340,336-349,351-362,369-378,374-38
 F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:342/Modified site: erythro-beta-hydroxyaspartic acid (Asn) #status experimental
 F:490,492/Binding site: chromotrolin sulfate (Ser) (covalent) (partial) #status experim

Query Match 8.7%; Score 227; DB 1; Length 575;
 Best Local Similarity 25.7%; Pred. No. 2.2e-06;
 Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

OY 16 PEPGGEHTARACCSAGACTSHHATMKRQAEEACILRGALSTVRAGAEELRAVIA 75
 DB 23 PEPGSGQCEVH-----CFALYGPATFELNQCIDGRLHLMVRSVADVISL 73
 OY 76 LTRAPGPGGSGKDLFWALE-----RRRSHCHLENPPLGFSWLSDDPGGLSDDT 127

DB 74 LTN---GDGGVGRRL-WIGLPLPGGDPKRLG-----PLRGFWYTGDNNTSYS--- 120
 OY 128 LOWEEPGRS-----CTARCAVLATGVEPAGKEKRECHLRANGYLCKYQFVCLP 180
 DB 121 -RMARLDNGAPLDCGLPCVAVSA--EATVSEPT-WEEOCEVKADGFLCEHFPAFPCR 176
 OY 181 -APRGAAS---NLSRAPFOLHSALDFSPTEVSALCRQPLSVYCIAD----- 229
 DB 177 PLAVEPGAAAASVITGTFFAANGAPFQALPVGSSAAV---APLGIQLMCTAPPAAVGG 233
 OY 230 ---ELIARD-----KLSGVLCPCP-----GRLRAG--KCAEL-- 259
 DB 234 HMAEAPGAMDCVENGCEHCACNAIFGAPRCOCPPAGALQADGRCTASATQSCNDLCE 293
 OY 260 ---PNCIDLDGFCACATGFEIKGRGRC 286
 DB 294 HFCVFN-PDQPGSVSCMCEGYLADQHC 323

RESULT 2

A:Reference number: A60501
 A:Molecule type: precursor - mouse
 N:Alternate names: fetomodulin
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
 A:Accession: S08488; A52001; A60501
 R.Dittman, W.A.; Majerus, P.W.
 Nucleic Acids Res. 17, 802, 1989
 A>Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted
 A:Reference number: S08488; MUID:89128454
 A:Accession: S08488
 A:Molecule type: mRNA
 A:Residues: 1-577 <DIT>
 A:Cross-references: EMBL:X14432; NID:954781; PID:CAA32597.1; PID:954782
 R.Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.
 J. Biol. Chem. 263, 15815-15822, 1988
 A>Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat
 A:Reference number: A32001; MUID:89008498
 A:Accession: A32001
 A:Molecule type: mRNA
 A:Residues: 97-577 <D12>
 A:Cross-references: GB:J04060
 R.Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.
 Dev. Biol. 140, 113-122, 1990
 A>Title: Identification of fetomodulin, a surface marker protein of fetal development
 A:Reference number: A60501; MUID:90292331
 A:Accession: A60501
 A:Molecule type: protein
 A:Residues: 19-22;330-343;479-489;545-555;562-575 <IMA>
 C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor
 F:24-165/Domain: C-type lectin homology <LCH>
 F:244-279/Domain: EGF homology <EG1>
 F:287-322/Domain: EGF homology <EG2>
 F:328-361/Domain: EGF homology <EG3>
 F:368-403/Domain: EGF homology <EG4>
 F:407-438/Domain: EGF homology <EG5>
 F:444-479/Domain: EGF homology <EG6>

Query Match 7.4%; Score 192; DB 2; Length 577;
 Best Local Similarity 24.7%; Pred. No. 6.4e-06;
 Matches 75; Conservative 34; Mismatches 125; Indels 70; Gaps 15;

OY 37 CYSLHHATMKRQAEEACILRGALSTVRAGAEELRAVIALLRAGPGGSGKDLFWAL 96
 DB 35 CFALYGPATFELNQCIDGRLHLMVRSVADVISL-----SQSSMDLGPWIGL 88
 OY 97 ERRRSHC--TLENEPLRGFSWLSDDPGGLESPDLOWEEPGRSCTARRAVLQATGCV-- 152
 DB 89 QLPQG--CDPEVHGLGPIRGFWYTGDN---HTYSRWARPNDQ--TAPLGLPCVTVSTAT 142

GenCore version 4.5.
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:33:30 ; Search time 13.46 Seconds

(without alignments)
1409.553 Million cell updates/sec

Title: US-09-902-713b-96

Perfect score: 2605

Sequence: 1 MRPAFALCLLMQALMPGPG.....LRDRAEGALLAESPLGSSDA 490

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 227 | 8.7 | 575 | TRBM_HUMAN | P07204 homo sapien |
| 2 | 192 | 7.4 | 577 | TRBM_MOUSE | P15306 mus musculu |
| 3 | 143.5 | 5.5 | 273 | MT75_MOUSE | Q9h9p2 homo sapien |
| 4 | 139.5 | 5.4 | 273 | MT75_MOUSE | Q9c8m0 mus musculu |
| 5 | 134.5 | 5.2 | 443 | FBL4_CRIGR | Q53058 cricetulus |
| 6 | 133.5 | 5.1 | 443 | FBL4_HUMAN | Q95367 homo sapien |
| 7 | 130.5 | 5.0 | 443 | FBL4_MOUSE | Q9w7j9 mus musculu |
| 8 | 128 | 4.9 | 592 | DLI3_MOUSE | Q88516 mus musculu |
| 9 | 126 | 4.8 | 589 | DLI3_RAT | Q88671 rattus norv |
| 10 | 126 | 4.8 | 3695 | LMA5_HUMAN | O15230 homo sapien |
| 11 | 124.5 | 4.8 | 2477 | FINC_RAT | P04337 rattus norv |
| 12 | 124 | 4.7 | 933 | PERT_HUMAN | P07202 homo sapien |
| 13 | 122 | 4.7 | 618 | DLI3_HUMAN | Q9w7j7 homo sapien |
| 14 | 119 | 4.6 | 646 | LEM3_BOVIN | P42201 bos sapien |
| 15 | 119 | 4.6 | 2871 | FBN1_BOVIN | Q91t36 sus scrofa |
| 16 | 119 | 4.6 | 2907 | FBN2_MOUSE | Q61555 mus musculu |
| 17 | 118.5 | 4.5 | 1408 | SERR_DROME | P18168 drosophila |
| 18 | 118.5 | 4.5 | 4655 | LRP2_HUMAN | P98164 homo sapien |
| 19 | 118 | 4.5 | 448 | FBL5_HUMAN | Q9uwx5 homo sapien |
| 20 | 117.5 | 4.5 | 448 | FBL5_MOUSE | Q9w7h9 mus musculu |
| 21 | 117.5 | 4.5 | 2318 | NTC3_MOUSE | Q61882 mus musculu |
| 22 | 117 | 4.5 | 2871 | FBN1_BOVIN | P98133 bos sapien |
| 23 | 117 | 4.5 | 2871 | FBN1_HUMAN | P35555 mus musculu |
| 24 | 117 | 4.5 | 2871 | FBN1_MOUSE | Q61554 mus musculu |
| 25 | 116.5 | 4.5 | 448 | FBL5_RAT | Q9w7h8 rattus norv |
| 26 | 116 | 4.5 | 2911 | FBN2_HUMAN | P35556 homo sapien |
| 27 | 115 | 4.4 | 174 | PAR3_MOUSE | Q09949 mus musculu |
| 28 | 115 | 4.4 | 1376 | SCAP_CRIGR | P97260 cricetulus |
| 29 | 115 | 4.4 | 4753 | LRP_CAEEL | Q04833 caenorhabd |
| 30 | 114.5 | 4.4 | 1456 | MANR_HUMAN | P22897 homo sapien |
| 31 | 114.5 | 4.4 | 1964 | NTC4_MOUSE | P31695 mus musculu |
| 32 | 113.5 | 4.4 | 676 | PRTS_HUMAN | P07225 homo sapien |
| 33 | 113.5 | 4.4 | 2254 | CCAG_RAT | O54898 rattus norv |

ALIGNMENTS

| RESULT | ID | TRBM_HUMAN | STANDARD | PRT | 575 AA |
|--------|-------|------------|----------|-----|------------|
| 34 | 113 | 4.3 | 2139 | 1 | CRB_DROME |
| 35 | 112.5 | 4.3 | 909 | 1 | LDLI_XENLA |
| 36 | 112.5 | 4.3 | 4660 | 1 | LRP2_RAT |
| 37 | 112 | 4.3 | 1221 | 1 | FBL2_MOUSE |
| 38 | 111 | 4.3 | 596 | 1 | MKT7_YEAST |
| 39 | 111 | 4.3 | 2437 | 1 | NOTC_BRARE |
| 40 | 110.5 | 4.2 | 1375 | 1 | NID2_HUMAN |
| 41 | 110 | 4.2 | 675 | 1 | PRTS_MOUSE |
| 42 | 110 | 4.2 | 776 | 1 | SNIL_RAT |
| 43 | 109.5 | 4.2 | 649 | 1 | PRTS_MACMU |
| 44 | 109.5 | 4.2 | 926 | 1 | PERT_PIG |
| 45 | 109 | 4.2 | 675 | 1 | PRTS_RAT |

| | | | | |
|----|--|----------|-----|--------|
| 1 | TRBM_HUMAN | STANDARD | PRT | 575 AA |
| AC | P07204 | | | |
| DT | 01-APR-1988 (Rel. 07, Created) | | | |
| DT | 01-FEB-1991 (Rel. 17, Last sequence update) | | | |
| DT | 01-MAR-2002 (Rel. 41, Last annotation update) | | | |
| DE | Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen). | | | |
| GN | THBD OR THRM. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88004395; PubMed=2820710; | | | |
| RA | Suzuki K., Kusumoto H., Devashiki Y., Nishioka J., Maruyama I., | | | |
| RA | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.; | | | |
| RT | "Structure and expression of human thrombomodulin, a thrombin | | | |
| RT | receptor on endothelium acting as a cofactor for protein C | | | |
| RT | activation." | | | |
| RL | EMBO J. 6:1891-1897(1987). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88024950; PubMed=2822087; | | | |
| RA | Men D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.; | | | |
| RT | "Human thrombomodulin: complete cDNA sequence and chromosome | | | |
| RT | localization of the gene." | | | |
| RL | Biochemistry 26:4350-4357(1987). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=87317655; PubMed=2819876; | | | |
| RA | Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.; | | | |
| RT | "Human thrombomodulin gene is intron depleted: nucleic acid sequences | | | |
| RT | of the cDNA and gene predict protein structure and suggest sites of | | | |
| RT | regulatory control." | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88227901; PubMed=2836377; | | | |
| RA | Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H., | | | |
| RA | Devashiki Y., Maruyama I., Suzuki K.; | | | |
| RT | "Gene structure of human thrombomodulin, a cofactor for thrombin- | | | |
| RT | catalyzed activation of protein C." | | | |
| RL | J. Biochem. 103:281-285(1988). | | | |
| RN | [5] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., | | | |
| RA | Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., | | | |
| RA | Bailey J., Barlow K.F., Bates K.N., Beare D.M., Brown A.J., | | | |
| RA | Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown M.P., | | | |
| RA | Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., | | | |
| RA | Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., | | | |
| RA | Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R., | | | |
| RA | Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M., | | | |
| RA | Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., | | | |

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levenshain M.H., Leverisha M., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
RA Milne S., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuse C.D., Smith M.I., Soderlund C., Steward C.A., Sulston J.E.,
RA Tracey A., Symons N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Whitehead S.L., Whitaker P., Willey D.E., Wall M., Wallis J.M.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RA "The DNA sequence and comparative analysis of human chromosome 20.",
RA Nature 414:865-871(2001).
RN [6]
RN CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS
RN MEDLINE-94029900; PubMed-8216207;
RN Gerlitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,
RN Grinnell B.W.,
RN "Identification of the predominant glycosaminoglycan-attachment site
RN in soluble recombinant human thrombomodulin: potential regulation of
RN functionality by glycosyltransferase competition for serine474.",
RN Biochem. J. 295:131-140(1993).
RN [7]
RN STRUCTURE BY NMR OF 389-407.
RN MEDLINE-96007474; PubMed-7559494;
RN Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.,
RN "The structure of a 19-residue fragment from the C-loop of the fourth
RN epidermal growth factor-like domain of thrombomodulin",
RN J. Biol. Chem. 270:23366-23372(1995).
RN [8]
RN STRUCTURE BY NMR OF 364-407.
RN MEDLINE-96100636; PubMed-8528067;
RN Meininger D.P., Hunter M.J., Komives E.A.,
RN "Synthesis, activity, and preliminary structure of the fourth
RN EGF-like domain of thrombomodulin",
RN Protein Sci. 4:1683-1695(1995).
RN [9]
RN STRUCTURE BY NMR OF 427-444.
RN MEDLINE-95034791; PubMed-7947766;
RN Srivatsan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.,
RN "Thrombin-bound structure of an EGF subdomain from human
RN thrombomodulin determined by transferred nuclear Overhauser
RN effects.",
RN Biochemistry 33:13553-13560(1994).
RN [10]
RN STRUCTURE BY NMR OF 427-444.
RN MEDLINE-96276211; PubMed-8745396;
RN Hrabal R., Komives E.A., Ni F.,
RN "Structural resiliency of an EGF-like subdomain bound to its target
RN protein thrombin.",
RN Protein Sci. 5:195-203(1996).
RN [11]
RN STRUCTURE BY NMR OF 405-444.
RN MEDLINE-98035729; PubMed-9367781;
RN Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.,
RN "Structure of the fifth EGF-like domain of thrombomodulin: an
RN EGF-like domain with a novel disulfide-bonding pattern.",
RN J. Mol. Biol. 273:913-926(1997).
RN [12]
RN VARIANT TED TYR-486.
RN MEDLINE-95111115; PubMed-7811989;
RN Oehlin A.-K., Marlar R.A.,
RN "The first mutation identified in the thrombomodulin gene in a
RN 45-year-old man presenting with thromboembolic disease",
RN Blood 85:330-336(1995).
RN [13]
RN VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.
RN MEDLINE-97341986; PubMed-9198186;
RN Oehlin A.-K., Norlund L., Marlar R.A.,
RN "Thrombomodulin gene variations and thromboembolic disease.",
RN Thromb. Haemost. 78:396-400(1997).
RN [14]
RN VARIANT VAL-473.
RN MEDLINE-97206519; PubMed-9157575;
RN Norlund L., Holm J., Zoller B., Oehlin A.-K.,
RN "A common thrombomodulin amino acid dimorphism is associated with
RN myocardial infarction.",
RN Thromb. Haemost. 77:248-251(1997).
RN [15]
RN VARIANT THR-43.
RN MEDLINE-99057299; PubMed-9843165;
RN Dodgen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,
RN Stubbs P.J., Manger Cats V., Ireland H.,
RN "A mutation in the thrombomodulin gene, 1276 to A coding for Ala25Thr,
RN and the risk of myocardial infarction in men.",
RN Thromb. Haemost. 80:743-748(1998).
RN [16]
RN VARIANT VAL-473.
RN MEDLINE-21143723; PubMed-11245641;
RN Wu K.K., Aleksić N., Ahn C., Boerwinkle E., Folsom A.R.,
RN Juneja H.,
RN "Thrombomodulin Ala45Val polymorphism and risk of coronary heart
RN disease.",
RN Circulation 103:1386-1389(2001).
RN [17]
RN FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
RN THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
RN COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
RN ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
RN SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
RN FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF
RN THROMBIN GENERATED.
RN SUBCELLULAR LOCATION: Type I membrane protein.
RN TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING
RN THROMBOMODULIN.
RN POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED
RN RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).
RN DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,
RN ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE
RN DEFECTS OF THE HAEMOSTATIC SYSTEM WHICH CREATES A TENDENCY TO THE
RN OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE
RN PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.
RN SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
RN DATABASE: NAME-PROV. NOTE-CD guide CD141 entry;
RN WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd141.htm".
RN [18]
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RN or send an email to license@isb-sib.ch).
RN EMBL; X03495; CAA29045.1;
RN EMBL; M16552; AAB59508.1;
RN EMBL; J02973; AAA61175.1;
RN EMBL; D00210; BAA00149.1;
RN EMBL; AL049651; CAB51954.1;
RN PIR; A27073; A27073.
RN PIR; A28307; A28307.
RN PIR; A28307; A28307.
RN PIR; A29680; A29680.
RN PIR; IEQT; 15-NOV-95.
RN PDB; 1FGD; 20-JUN-96.
RN PDB; 1TMR; 08-JUN-95.
RN PDB; 1ZAG; 29-JAN-96.
RN PDB; 1ADK; 24-DEC-97.
RN PDB; 2ADK; 24-DEC-97.
RN GlycoSiteDB; P07204;
RN MIM; 188040;
RN InterPro; IPR000152; Asx_hydroxyl.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:33:56 ; Search time 30.72 Seconds

(without alignments)
2759.359 Million cell updates/sec

Title: US-09-902-713b-96

Perfect score: 2605
Sequence: 1 MRPAFALCLIMQALWPGPG.....LRDRAEGALLAESPLGSSDA 490

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP.TREMBL.19: **
2: SP.Bacteria: **
3: SP.Fungi: **
4: SP.Human: **
5: SP.Invertebrate: **
6: SP.Mammal: **
7: SP.Mhc: **
8: SP.Organelle: **
9: SP.Phage: **
10: SP.Plant: **
11: SP.Protoct: **
12: SP.Virus: **
13: SP.Vertebrate: **
14: SP.Unclassified: **
15: SP.Virus: **
16: SP.Bacteriaph: **
17: SP.Archeap: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|--------------------|
| 1 | 1586.5 | 60.9 | 459 | 11 Q9D624 | Q9D624 mus musculu |
| 2 | 1584.5 | 60.8 | 459 | 11 Q9CXA8 | Q9CXA8 mus musculu |
| 3 | 1574.5 | 60.4 | 459 | 11 Q9DC55 | Q9DC55 mus musculu |
| 4 | 957.5 | 36.8 | 212 | 4 Q9P096 | Q9P096 homo sapien |
| 5 | 350 | 13.4 | 757 | 4 Q9HC00 | Q9HC00 homo sapien |
| 6 | 336.5 | 12.9 | 765 | 11 Q91ZV1 | Q91ZV1 mus musculu |
| 7 | 336.5 | 12.9 | 765 | 11 Q91V98 | Q91V98 mus musculu |
| 8 | 309 | 11.9 | 652 | 4 Q9NPY3 | Q9NPY3 homo sapien |
| 9 | 307 | 11.8 | 652 | 4 Q90274 | Q90274 homo sapien |
| 10 | 305 | 11.7 | 643 | 11 Q90126 | Q90126 homo sapien |
| 11 | 303 | 11.6 | 644 | 11 Q89103 | Q89103 mus musculu |
| 12 | 301 | 11.6 | 643 | 11 Q9ET61 | Q9ET61 rattus norv |
| 13 | 227 | 8.7 | 468 | 4 Q9UC32 | Q9UC32 homo sapien |
| 14 | 193 | 7.4 | 577 | 11 Q93370 | Q93370 rattus norv |
| 15 | 143.5 | 5.5 | 184 | 4 Q9HCY3 | Q9HCY3 homo sapien |
| 16 | 142 | 5.5 | 955 | 4 Q96DN2 | Q96DN2 homo sapien |

| | | | | | |
|----|-------|-----|------|-----------|--------------------|
| 17 | 139.5 | 5.4 | 273 | 11 Q9CXM0 | Q9CXM0 mus musculu |
| 18 | 139 | 5.3 | 461 | 11 P97883 | P97883 rattus norv |
| 19 | 139 | 5.3 | 1394 | 5 Q9V589 | Q9V589 drosophila |
| 20 | 138.5 | 5.3 | 443 | 4 Q9H3D5 | Q9H3D5 homo sapien |
| 21 | 137 | 5.3 | 1637 | 6 Q9XSV8 | Q9XSV8 bos taurus |
| 22 | 135.5 | 5.2 | 624 | 4 Q96JG5 | Q96JG5 homo sapien |
| 23 | 133.5 | 5.1 | 443 | 4 Q96TF5 | Q96TF5 homo sapien |
| 24 | 132.5 | 5.1 | 443 | 11 Q9J0M6 | Q9J0M6 mus musculu |
| 25 | 132.5 | 5.1 | 1238 | 4 Q9UEP9 | Q9UEP9 homo sapien |
| 26 | 132.5 | 5.1 | 1238 | 4 Q9Y6Q0 | Q9Y6Q0 homo sapien |
| 27 | 132.5 | 5.1 | 1238 | 4 Q9UNK8 | Q9UNK8 homo sapien |
| 28 | 132 | 5.1 | 1233 | 4 Q9UE17 | Q9UE17 homo sapien |
| 29 | 132 | 5.1 | 3507 | 5 Q23587 | Q23587 caenorhabd |
| 30 | 131.5 | 5.0 | 1238 | 4 Q9Y219 | Q9Y219 homo sapien |
| 31 | 130.5 | 5.0 | 985 | 12 Q67643 | Q67643 gallid hep |
| 32 | 129.5 | 4.9 | 1106 | 5 Q17494 | Q17494 caenorhabd |
| 33 | 128.5 | 4.9 | 803 | 6 Q9GWT9 | Q9GWT9 macaca fasc |
| 34 | 128.5 | 4.9 | 1001 | 4 Q9BYH8 | Q9BYH8 homo sapien |
| 35 | 128.5 | 4.9 | 1404 | 5 Q9VB65 | Q9VB65 drosophila |
| 36 | 124.5 | 4.8 | 2809 | 4 Q96UP8 | Q96UP8 homo sapien |
| 37 | 123 | 4.7 | 1260 | 4 Q9Y6P9 | Q9Y6P9 homo sapien |
| 38 | 122.5 | 4.7 | 374 | 4 Q96NF3 | Q96NF3 homo sapien |
| 39 | 122.5 | 4.7 | 374 | 4 Q96NC5 | Q96NC5 homo sapien |
| 40 | 122.5 | 4.7 | 1479 | 11 Q64449 | Q64449 mus musculu |
| 41 | 122.5 | 4.7 | 2319 | 11 Q9R172 | Q9R172 rattus norv |
| 42 | 122 | 4.7 | 248 | 11 Q9Z211 | Q9Z211 mus musculu |
| 43 | 122 | 4.7 | 1963 | 6 Q28019 | Q28019 bos taurus |
| 44 | 122 | 4.7 | 2471 | 11 Q9QW30 | Q9QW30 rattus sp., |
| 45 | 121.5 | 4.7 | 374 | 11 Q9Z209 | Q9Z209 cricetus |

ALIGNMENTS

| RESULT | 1 | PRELIMINARY: | PRT: | 459 AA. |
|--------|---|--------------|------|---------|
| Q9D624 | Q9D624 | | | |
| AC | 01-JUN-2001 (TREMBLrel. 17, Created) | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) | | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) | | | |
| DE | 1200003C23RIK PROTEIN. | | | |
| GN | 1200003C23RIK. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=C57BL/6J; TISSUE=HEAD; | | | |
| RC | MEDLINE=21085660; Pubmed=11217851; | | | |
| RA | Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., | | | |
| RA | Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., | | | |
| RA | Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S., | | | |
| RA | Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R., | | | |
| RA | Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., | | | |
| RA | Fleischmann M., Gaasterland T., Giesi C., King B., Kochiwa H., | | | |
| RA | Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., | | | |
| RA | Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., | | | |
| RA | Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G., | | | |
| RA | Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F., | | | |
| RA | Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., | | | |
| RA | Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H., | | | |
| RA | Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P., | | | |
| RA | Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., | | | |
| RA | Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., | | | |
| RA | Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., | | | |
| RA | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., | | | |
| RA | Hashizaki Y., | | | |
| RT | *Functional annotation of a full-length mouse cDNA collection.*; | | | |
| RT | Nature 409:685-690(2001). | | | |
| EMBL | AK014681; BAB2502.1; - | | | |
| DR | HSSP; P00743; ICCF. | | | |

| QY | 1 | MR | FA | FL | CL | LM | AL | MP | PG | GG | EH | PA | DA | GS | SA | GA | CS | SL | HA | TM | KO | AA | EA | CI | LG | GA | 60 | | |
|----|-----|----|----|----|----|----|----|----|----|---|----|----|----|---|----|----|---|---|----|----|---|---|---|---|----------------------------|----|----|----|----|
| QY | 1 | MR | FA | FL | CL | LM | AL | MP | PG | GG <td>EH</td> <td>PA</td> <td>DA</td> <td>GS <td>SA</td> <td>GA</td> <td>CS <td>SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td></td></td></td></td></td></td> | EH | PA | DA | GS <td>SA</td> <td>GA</td> <td>CS <td>SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td></td></td></td></td></td> | SA | GA | CS <td>SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td></td></td></td></td> | SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td></td></td></td> | HA | TM | KO <td>AA <td>EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td></td></td> | AA <td>EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td></td> | EA <td>CI <td>LG <td>GA</td> <td>60</td> </td></td> | CI <td>LG <td>GA</td> <td>60</td> </td> | LG <td>GA</td> <td>60</td> | GA | 60 | | |
| Db | 1 | MR | FA | FL | CL | LM | AL | MP | PG | GG | EH | PA | DA | GS <td>SA</td> <td>GA</td> <td>CS <td>SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI</td> <td>LG <td>GA</td> <td>60</td> </td></td></td></td></td></td> | SA | GA | CS <td>SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI</td> <td>LG <td>GA</td> <td>60</td> </td></td></td></td></td> | SL <td>HA</td> <td>TM</td> <td>KO <td>AA <td>EA <td>CI</td> <td>LG <td>GA</td> <td>60</td> </td></td></td></td> | HA | TM | KO <td>AA <td>EA <td>CI</td> <td>LG <td>GA</td> <td>60</td> </td></td></td> | AA <td>EA <td>CI</td> <td>LG <td>GA</td> <td>60</td> </td></td> | EA <td>CI</td> <td>LG <td>GA</td> <td>60</td> </td> | CI | LG <td>GA</td> <td>60</td> | GA | 60 | | |
| QY | 61 | LS | TV | RG | AG | LA | RV | AL | LL | LR | AG | PG | GG | SK | DL | FW | AL | ER | RR | SH | CT | LE | NE | PL | RG | SW | LS | DP | 12 |
| Db | 61 | LS | TV | RG | AG | LA | RV | AL | LL | LR | AG | PG | GG | SK | DL | FW | AL | ER | RR | SH | CT | LE | NE | PL | RG | SW | LS | DP | 12 |
| QY | 121 | GG | LE | ST | LT | OW | EE | PP | OR | ST | AR | CA | VI | AT | GG | VE | PA | GG | KE | KH | CH | RA | NG | LY | CT | OE | EV | LC | 18 |
| Db | 121 | GG | LE | ST | LT | OW | EE | PP | OR | ST | AR | CA | VI | AT | GG | VE | PA | GG | KE | KH | CH | RA | NG | LY | CT | OE | EV | LC | 18 |
| QY | 121 | ED | SE | OP | LP | W | RE | PP | OR | ST | AR | CA | VI | AT | GG | VE | PA | GG | KE | KH | CH | RA | NG | LY | CT | OE | EV | LC | 18 |
| Db | 121 | ED | SE | OP | LP | W | RE | PP | OR | ST | AR | CA | VI | AT | GG | VE | PA | GG | KE | KH | CH | RA | NG | LY | CT | OE | EV | LC | 18 |

| | | | | |
|----|-----|------------------------------------|----------------------------|-----|
| QY | 181 | ARRPGASNLSTRAPFOLHSALDSEPPGEVSA | LMCGQPLISVITIADEIGARMOKL-S | 23 |
| | | | | |
| Db | 181 | ARRPGASNLSTQAPFRSLSSALDFSPPGREVSAM | CGDLSVSTCJOEETSAHMDGLFP | 24 |
| | | | | |
| QY | 240 | GDVLPCPCGRTLRNGKAEINCLDLDGGRACACNG | FELGKRGKSCVYSGSGQPLTGGT | 29 |
| | | | | |
| Db | 241 | GVVLPCGGRILNKGKVELPDCILHGTCTECQAV | FELGKRGKSCFTRKVEOLTLTGGT | 30 |
| | | | | |
| QY | 300 | GYPTRRPPATLTSVPOPTWPIRVDKLTGTPLP | VEDNSVYSIPEIRPMGSGOSTMSTLO | 35 |
| | | | | |
| Db | 301 | KLPETNNVATPAGAVTNRKTPMGQVYDKR | GEMP-----QVEIILMGQSTLPIITQ | 35 |
| | | | | |
| QY | 360 | MSLQAESKATITPBGSGVYSIKFNSTTSATQAD | DSSAVVFLFVSTAVVYVILMTVLG | 41 |
| | | | | |
| Db | 351 | KTPQKPKRYICTPBGSVV-LNTTSSPPVSL | FFDTISVTFVLSIAIVIVLVIITVLG | 408 |
| | | | | |
| QY | 420 | LVKLCFHEBSPQKRKESMGPGLIEDP | EPFALGSSASACITNNKGVG | 467 |
| | | | | |
| Db | 409 | LEKLCFHKRSRSPRGKALDSPGVEDCA | DEATSLHHSSTQCITDGVKSG | 456 |
| | | | | |

| AC | ID | ORGXAB | PRELIMINARY; | PRT; | 459 | AA. |
|---|---|--------|--------------|------|-----|-----|
| 09CXAB | 09CXAB | | | | | |
| 01-JUN-2001 | (Tremblrel, 17, Created) | | | | | |
| 01-JUN-2001 | (Tremblrel, 17, Last sequence update) | | | | | |
| 01-DEC-2001 | (Tremblrel, 19, Last annotation update) | | | | | |
| 1200003C23R1K | PROTEIN. | | | | | |
| 1200003C23R1K | | | | | | |
| Mus musculus | (Mouse) | | | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | |
| Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | | | |
| NCBI_Taxid=10090; | | | | | | |
| [1] | | | | | | |
| SEQUENCE FROM N.A.A. | | | | | | |
| STRAIN=C57BL/6J; TISSUE=EMBRYONIC LUNG; | | | | | | |
| MEDLINE=21085660; PubMed=11217851; | | | | | | |
| Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., | | | | | | |
| Akawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., | | | | | | |
| Alzawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I., | | | | | | |
| Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R., | | | | | | |
| Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., | | | | | | |
| Pleischmann W., Gasterland T., Glissl C., King B., Koehne H., | | | | | | |
| Kuehl J., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., | | | | | | |

RA Schriber L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.
RA Sakai K., Okido T., Futuro M., Aono H., Balzarelli R., Barsh G.,
RA Blake U., Boiffelli D., Bojunga N., Carninci P., de Donato M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombers P.,
RA Nordone P., Ring B., Ringwald N., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibaeva T., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Wiltaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
DR EMBL: AK018432; BAB31209.1; -
DR HSSP, P00743; ICF.
DR MGD; MG1:1914114; 1200003C23RLK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000514; EGF-like.
DR InterPro: IPR001304; EGF-like.
DR SMART; SM00034; C1ECT; 1.
DR SMART; SM00081; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00041; C_Type_Lectin_2; 1.
SQ SQUENCE 455 AA; 49150 MW; A649730F15122CD CRC64;

| | | | | |
|-----------------------|-----------------|------------------|-----------|------------|
| Query Match | 60.8% | Score 1584.5 | DB 11 | Length 459 |
| Best Local Similarity | 66.0% | Pred. No. 8e-122 | | |
| Matches 309 | Conservative 37 | Mismatches 109 | Indels 13 | Gaps 3 |

```

0Y 1 MRAPALCLLNALMPGGCGGHPADPAGCSASGACYSLHHATMKRQAAEACILRGA 60
Db 1 MRALALCLLCPAFMPGNGEHPADPAPARCASAGACYSLHHATFKRAAEACILRGT 60
0Y 61 LSTVRAGALRALVALLLRAGCGPGGSKDILFWALERRSHCTLENELRGSWMLSSDP 120
Db 61 LSTVHSGSEFQAQVLLLRAGPGPGGSKDILFWALERSISQCTQKEEPLRGSWMLHPS 120
0Y 121 GGLJEDSTLQWVEEPRSTARCACVLAQTGVEPAGNKEMRCHLRANGYLCKYQFVLCR 180
Db 121 EDESESPLEWVEEPRSTCKKCAALQATRVBEVAGNKEMRCHLRIDGYLCKYQFVLCR 180
0Y 181 ABRPGAASNLSPARFOLSHLSALDFSPGTEVSALCGOLPISVTCILADEIGARDKL-S 239
Db 181 ABRPGAASNLSPQAFRLSSALDFSPGTIEVSAMCPGDLSSVSTIQEETSLHMDGLFR 240
0Y 240 GNYLCPCGCRGLRAGCAELPNCILDDIGGFACRCANGFELGKDGRCVYSGEGOPFLIGT 299
Db 241 GNYLCPCGCRGLRAGCAELPNCILDDIGDTECCAVGFEIGKDGRCSECTKVEEQLTBOG 300
0Y 300 GYPTRRPRPATLSPVPORTMPIRYDEKLGETPLVPRQDMSVYSIPIPRMWSQSTJMSLTQ 359
Db 301 KLPTENVATATPAGAVNTRMPGQVYDKPRGMP-----QVETLLQMGQSTLPTIQ 350
0Y 360 MSLOAESKATITPBGSVYSIKFSNSTSAAAPQAFDSSAVVFLFVSAAVVALVILMTFYAG 419
Db 351 KTPQTKRPVYGTBPGSVY--LNTYSSPVSLFEDTSSVFLVILSVLAVIVLVLLITVLG 408
0Y 420 LVKLCFHEPSSQBRKESMCPGLEDPREPALIGSSASHACTNNKYVG 467
Db 409 LFKLCFHKRSSRKIGKALDPSVECDAAERTSLHSHSITQCTIDIGYKSG 456

```

| | | |
|--------|-------------------------|--|
| RESULT | 3 | |
| Q9DC55 | | |
| ID | Q9DC55 | |
| AC | Q9DC55 | |
| DT | 01-JUN-2001 | (TEMBLrel. 17, Created) |
| DT | 01-JUN-2001 | (TEMBLrel. 17, Last sequence update) |
| DT | 01-DEC-2001 | (TEMBLrel. 19, Last annotation update) |
| DE | 12000003C23RIK PROTEIN. | |
| DE | 12000003C23RIK | |

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=LUNG;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aachii J., Fukuda S.,
 RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuhl P., Lewis S., Matsuo Y., Nkaido I., Pesole G., Quackenbush J.,
 RA Schirini L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bono M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Jordon P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyrshew-Boris A., Yoshida K., Hasegawa T., Kawaji H., Kohlsunk S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection";
 RT Nature 409:685-690(2001).
 RL EMBL: AK004557; BAB3370.1; -
 DR HSSP: P00743; ICCF
 DR MGP: MGI:1914114; 1200003C23Rik.
 DR InterPro: IPR000152; Asx_Hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_C.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00010; ASX_Hydroxyl; UNKNOWN_1.
 DR PROSITE: PS00041; C_Type_Lectin_2; 1.
 DR PROSITE: PS50041; C_Type_Lectin_2; 1.
 SQ SEQUENCE 459 AA; 49053 MW; 6CB25D048A56A7B CRC64;

Query Match 60.4%; Score 1574.5; DB 11; Length 459;
 Best Local Similarity 65.4%; Pred. No. 5.3e-121;
 Matches 306; Conservative 39; Mismatches 110; Indels 13; Gaps 3;

QY 1 MRPALCLCLMOLMPGGGGEPTRADRAGSAGACTSLHATMKROAEACILRGGA 60
 DB 1 MRPALCLCLCAFPMPRONGEHPTRADRACASAGACTSLHASFRRRAEEACSLRGGT 60
 QY 61 LSTVBAELRAVLALLRAGPGGGSKDLFLWALERRRSHCTLENEPLRGFWLSLSDP 120
 DB 61 LSTVHSGSDPQVALLLRAGPGGGSKDLFLWALERSISCTOEKPELRGFWLSLHPS 120
 QY 121 GLESDTLQWVEPDRSTARCANVLOATGYEPAGKMKRCHLRANGILCYOFEVLCP 180
 DB 121 EDSESPSPWVEEPDRSTARCANVLOATGYEPAGKMKRCHLRANGILCYOFEVLCP 180
 QY 181 APRPAASNLSTYRAFPOLHSALDPSPTCEVSALCRGOLPISVCIADIEAICARMDKL-S 239
 DB 181 APRPAASNLSTYRAFPOLHSALDPSPTCEVSALCRGOLPISVCIADIEAICARMDKL-S 239
 QY 240 GDLVLCPCRGVYRAGKCAELPNCIDLDLGGFACCATGELGKGRSCVTSGEQPTLGGT 299
 DB 240 GDLVLCPCRGVYRAGKCAELPNCIDLDLGGFACCATGELGKGRSCVTSGEQPTLGGT 299
 QY 300 GVTERRPPATATSPVQPTWPIRVDEKIGETPLVPEQONSVSISIEIPRWSQSTMTIQ 359
 DB 300 KLPITNVTATTPAGATNKRTPQGVYDKDPEMP-----QVTEILLQWGSTLPTIQ 350
 QY 360 MSIAESKATITPGSGSVISKFNSTSSATPOAFDSSSAVFIFFVTAVVVLLTMTVLG 419
 DB 360 KLPITNVTATTPAGATNKRTPQGVYDKDPEMP-----QVTEILLQWGSTLPTIQ 350
 QY 351 KTPQRTKPKVGTGTPSGSV--LNTYSSPVSLLFDICSTVVFILVAVLVVLTITVLG 408

QY 420 LYKLCFHESPPSOPRKESMGPEGLSDPEPALGSSAHCTNNGVYG 467
 DB 409 LFLKCFHKRSRRTKGLALDPSGVECDAAEATSLHSSSTOCTDIGVXSG 456

RESULT 4
 ID Q9P096 PRELIMINARY; PRT: 212 AA.
 AC Q9P096;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE HSPC285 (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RA Ye M., Zhang O.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
 RT "Human partial CBS from c34+ stem cells";
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF161403; AAF28963.1; -
 FT NON_TER 1
 SQ SEQUENCE 212 AA; 22078 MW; 12E1194E7A8F75D1 CRC64;

Query Match 36.8%; Score 957.5; DB 4; Length 212;
 Best Local Similarity 81.5%; Pred. No. 8.2e-71;
 Matches 194; Conservative 5; Mismatches 10; Indels 29; Gaps 4;

QY 209 GREVSALCRGOLPISVCIADIEAICARMDLSDGVLCPCRGVYRAGKCAELPNCIDLDLGG 268
 DB 1 GINSRAMC-----CYPD-----PRTYRAGKCAELPNCIDLDLGG 34
 QY 269 FACECATGELKDKGR--SCVTSGEQPTLGGTGPTRRPPATATSPVQPTWPIRVDEK 326
 DB 35 FAANVSTGFVAG--EGRSACSVTSGEQPTLGGTGPTRRPPATATSPVQPTWPIRVDEK 93
 QY 327 LGETPLVPEQONSVSISIEIPRWSQSTMTLOMSIQAESKATITPGSGSVISKFNSTSS 386
 DB 94 LGETPLVPEQONSVSISIEIPRWSQSTMTLOMSIQAESKATITPGSGSVISKFNSTSS 153

QY 387 ATPQARDSSAVVEIPVSTAVVVLVLTMTVGLVLCFHESPPSOPRKESMGPEGL 444
 DB 154 ATPQARDSSAVVEIPVSTAVVVLVLTMTVGLVLCFHESPPSOPRKESMGPEGL 211

RESULT 5
 ID Q9HCU0 PRELIMINARY; PRT: 757 AA.
 AC Q9HCU0;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE TUMOR ENDOTHELIAL MARKER 1 PRECURSOR (ENDOSTALIN PROTEIN).
 GN TEM1 OR ENDOSTALIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20407466; PubMed=10941988;
 RX St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
 RA Montgomery E., Lal A., Riggin G.J., Lengauer C., Vogelstein B.,
 RA Kinzler K.W.;
 RT "Genes expressed in human tumor endothelium";
 RL Science 289:1197-1202(2000).
 [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE-21269274; PubMed-11084048;
 RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
 RT Garin-Chesa P., Park J.E., Relling W.J., Lenter M.C.,
 RT "Molecular cloning and characterization of Endostalin, a C-type
 lectin-like cell surface receptor of tumor Endothelium."
 RL J. Biol. Chem. 276:7408-7414(2001).
 DR EMBL; AF279142; AAG00867.1;
 DR EMBL; AJ295846; CAC34381.1;
 DR HSSP; P07204; IZAO-;
 DR Interpro: IPR000152; Asx_hydroxyl.
 DR Interpro: IPR000561; EGF-like.
 DR Interpro: IPR001881; EGF_Ca.
 DR Interpro: IPR001304; lectin_c.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00059; lectin_c.1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00179; EGF_Ca; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 KM Signal.
 FT SIGNAL.
 FT CHAIN.
 SQ SEQUENCE 757 AA; 80858 MW; C96363EAFBFFA0 CRC64;

Query Match 13.4%; Score 350; DB 4; Length 757;
 Best Local Similarity 26.1%; Pred. No. 2; 8e-20;
 Matches 122; Conservative 44; Mismatches 147; Indels 154; Gaps 20;

QY 7 LCLIMQALMPGGGGHPTA--DAGCSAGACYSLHHAATKROAEACILRGALSTV 64
 DB 5 LLLMAA--AGPTGQDPMAEPRAACPS--SCYALPPRRRTFLEAWRACREIGDLATPT 61
 QY 65 RAGALRAVALALRAGPGGGSKDLFWALERRSHCTLENPLRGFWLSSDGGLE 124
 DB 62 RTPEAQRVDSLVGAPA-----SRLMTIGLQROAQCO--RPLNGFTWTGD----- 109
 QY 125 SDT--LOWVEPQRS--CTARCAVLAQTGVEPAGKEMKCHLRANGYLCKYFYLCPA 181
 DB 110 QDTAFWMAQPGAGCPAPARCALEASG---EHRMLESGCTLAVGYLQGFEGECACRA 166
 QY 182 --PRPGAASNIYRAPFOLHSALDFSPGTEVSALCRGLPISVCIAD--ELGANW--- 235
 DB 167 LDPAAGAGAVYTPPHLVSTFEWLPFGSAVAOCQAGRASLLCVKQPEGVGWSRA 226
 DB 236 -----DKLSGVDLCPGPRYLIRAG-----KCAELP----- 260
 DB 227 GPLCLGTGSGPDNGGCEHECEVEVDGHSQRCETGFRILADGSCDPAQAPCEQCEP 286
 QY 261 -----NCLDDLGGFACCATGFEIGRD 282
 DB 287 GGPQGSCHCRIGFRAEDDPHRCVDTDCQIAGVCOQCMVYGGFEGYCEGHELEAD 346
 QY 283 GRSCTVSGE-----GQPTL-----GGTQVP-----TRRPAT 309
 DB 347 GISCSFAGAMGAQASODLDELDDSEDEDEDEAMKAFNGWTEPGLIMEPTQPDF 406
 QY 310 AFS-----PVQRTWPRIKDEKIGETPLVPEQDNSVTSI 343
 DB 407 ALAYRSPEDREPOLPYPEPTWPPPL-----SABRVP--YHSSVSLV 447

RESULT 6
 0912V1 PRELIMINARY; PRT; 765 AA.
 AC 0912V1.
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ENOCTALIN.
 GN TEM1.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWISS WEBSTER; TISSUE=THYMUS;
 RX MEDLINE-21486432; PubMed-11489895;
 RA Opavsky R., Havlenik P., Jurkovicova D., Garin M.T., Copeland N.G.,
 RA Gilbert D.J., Jenkins N.A., Bies J., Garfield S., Pastorekova S.,
 RA Que A., Wolff L.,
 RT "Molecular Characterization of the Mouse Tem1/endostalin Gene
 RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in
 RT Vivo."
 RL J. Biol. Chem. 276:38795-38807(2001).
 DR EMBL; AF388572; AAK84664.1;
 SQ SEQUENCE 765 AA; 81823 MW; 07B3225EDD13A03 CRC64;

Query Match 12.9%; Score 336.5; DB 11; Length 765;
 Best Local Similarity 32.8%; Pred. No. 3; 6e-19;
 Matches 94; Conservative 33; Mismatches 133; Indels 27; Gaps 12;

QY 7 LCLIMQALMPGGGGHPTADAGCSAGACYSLHHAATKROAEACILRGALSTVRA 66
 DB 5 LLLMAA--AGPTGQDPMAEPRAACPS--SCYALPPRRRTFLEAWRACREIGDLATPT 63
 QY 67 GAEALRAVALALRAGPGGGSKDLFWALERRSHCTLENPLRGFWLSSDGGLESD 126
 DB 64 PEAQRVDSLVGAPG-----LMTIGLQROAQCO--PORPLNGFTWTGD-----QD 111
 QY 127 T--LOWVE--EPQRSTARCAVLAQTGVEPAGKEMKCHLRANGYLCKYFYLCPA-- 181
 DB 112 TAFWMAQPGAGCPAPARCALEASG---EHRMLESGCTLAVGYLQGFEGECACRA 168
 QY 182 PRPGAASNIYRAPFOLHSALDFSPGTEVSALCRGLPISVCIADET--GANWKLKG 240
 DB 169 LEVGAAGAVYTPPHLVSTFEWLPFGSAVAOCQAGRASLLCVKQPEGVGWSOTG- 227
 QY 241 DYLCPCPPRYLRAGKCAELPNCDDL--GGFACCATGFEIGKDRSC 286
 DB 228 PLCPGTGCGPDNGGCEH--ECVEVDGAVSCRCSEGRILADGHS 271

RESULT 7
 091V98 PRELIMINARY; PRT; 765 AA.
 AC 091V98.
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE TUMOR ENDOTHELIAL MARKER 1 PRECURSOR (ENDOSTALIN).
 GN TEM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129S6/SVETAC; TISSUE=SPLEEN;
 RX MEDLINE-21486432; PubMed-11489895;
 RA Opavsky R., Havlenik P., Jurkovicova D., Garin M.T., Copeland N.G.,
 RA Gilbert D.J., Jenkins N.A., Bies J., Garfield S., Pastorekova S.,
 RA Que A., Wolff L.,
 RT "Molecular Characterization of the Mouse Tem1/endostalin Gene
 RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in
 RT Vivo."
 RL J. Biol. Chem. 276:38795-38807(2001).
 DR EMBL; AF388572; AAK84664.1;
 SQ SEQUENCE 765 AA; 81823 MW; 07B3225EDD13A03 CRC64;

RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in
 RT Vivo.
 RL J. Biol. Chem. 276:38795-38807(2001).
 DR EMBL: AF378758; AAL11995.1;
 DR EMBL: AF388573; AAK8465.1;
 RT Signal.
 FT CHAIN 1 17 POTENTIAL.
 FT SIGNAL 18 765 TUMOR ENDOTHELIAL MARKER 1.
 SQ SEQUENCE 765 AA; 81813 MM; 572A06CC15BC8C8D CRC64;

Query Match 12.9%; Score 336.5; DB 11; Length 765;
 Best Local Similarity 32.8%; Pred. No. 3.6e-19;
 Matches 94; Conservative 33; Mismatches 133; Indels 27; Gaps 12;

QY 7 LCLLQALMPGCGGHEPTADRGCSAGACYSLHNTMKROAEACILRGALSTVRA 66
 DB 5 LLLAVAAVPAALGOVPMTEPRACGSPS-SCYALFPRRTFLAMACELGSLATPRT 63
 QY 67 GAEFLAVLALRAGPGGGSKDLFWALERRSHCTLENEPLRGFSWLSDDPGLESD 126
 DB 64 PEARAVRSLVGVGANG-----LLMIGLQROARCCQ-POBPLRGFIWTTGD-----OD 111
 QY 127 T-LQWVE-EPQRCSTARCAVLQATGVEPAGMKEMRCHLANGYLCKYQPEVLCPA-- 181
 DB 112 TAFTHAQATGCPCPACACALDENSQ---EHRMLEGSCSTLAVDGLCPGEGACPALP 168
 QY 182 PRGGAASNLSTYRAPPOLHSAALDFSPPTGVSALRGQPLISVTCIADBI-GARMDKLSG 240
 DB 169 LEVGQAGPAVYTTTFFNLVSEFEMLPFQSVAAVQCGAGSALICVKQSGGWSQGTG- 227
 QY 241 DVLCPGPGRYLAKGCAELPCLDLI-GGFACECATGFLGKDGNSC 286
 DB 228 -PLCGTGGCGPNGGCEH--ECVEEVDGAVSCRCSEGFRLADGHSIC 271

RESULT 8
 ID 09NPY3 PRELIMINARY; PRT; 652 AA.
 AC 09NPY3;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE DJ737E23.1 (COMPLEMENT COMPONENT C1Q RECEPTOR).
 GN DJ737E23.1.
 OS 'Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sehra H.;
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL118508; CAC00597.1;
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR001304; Iectin_C.
 DR Pfam: PF00008; EGF_5.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00179; EGF_CA_3.
 DR SMART: SM00001; EGF_Like_2.
 DR PROSITE: PS00010; ASX_HYDROXYL_3.
 DR PROSITE: PS00041; C_TYPE_LLECTIN_2_1.
 DR PROSITE: PS01186; EGF_2_3.
 DR PROSITE: PS01187; EGF_CA_3.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
 KW Receptor; Repeat.
 SQ SEQUENCE 652 AA; 68560 MM; EECADPEAC55FCAC2 CRC64;

Query Match 11.9%; Score 309; DB 4; Length 652;

Best Local Similarity 21.4%; Pred. No. 5.3e-17;
 Matches 139; Conservative 82; Mismatches 205; Indels 224; Gaps 28;

QY 1 MRPAFLCLLQALMPGCGGHEPTADRGCSAGACYSLHNTMKROAEACILRGGA 60
 DB 1 MATSKLILLILLILLIQPGAGTADTEAYVCGT-ACVTAHSKILSAEAEQNHONCGN 59
 QY 61 LSTVRAGAE--LRAVLA--LLRAGPGGGSKDLFWALERRSHCTLENEPLRGFSW 115
 DB 60 LATVKSKEEAQHVQVLAQLLRREALTLARMSK---FWIGLQREKCKCLDPSLPLGFSW 116
 QY 116 LSSDPGLESIDLQWVEPQRCSTARCA--VLQATGVEPA--GWKEKRC-----HL 164
 DB 117 VG--GGEDTPYSNMWKEKLNCSIKRCVSLDLISQPLPLPRLPWSRSGPSPSPS 173
 QY 165 RANGYCKTQFVLCAPRGAASNLSTYRAPPOLHSAALDFSPPTGVS- 214
 DB 174 NIEGFVCKSPFKMCRPLALGGRGYTTTTPFTTSSLEAVFAANAVACGEGKDET 233
 QY 215 ----LCRGOLP-----ISVTCIADIGARD-----RL 238
 DB 234 QSHYFLCKEKAPDVFDMSSGPLCVSPKVCNPNNGCHODCFEGDGSFLCGCRPGFRL 293
 QY 239 SGDYL-----CCQPRY-----LRGKCAELP 260
 DB 294 LDDLVTGASRNPCSSSPRCGATCYLGPHGKNYTCRCPQGYQLDSSQLDCVDVDECQDSP 353
 QY 261 ---NCILDDLGFAECATGFE----- 278
 DB 354 CAQECVNTQGGFCEGVMGEPGPGPEGACQDVDECALSRPAQCTNIDGSFHCSEE 413
 QY 279 ----LCKDGRSC---VTSGEQP-----TLGCTGVTFRPPAT 309
 DB 414 GYVLAGEDEITQCCDQVDECVPGGGLCDSLCFNTQGSFHCGLPGWLAIPAGVSCITMGPSV 473
 QY 310 AHSVPQRTWPIRVD--EKIGETPIVPEQDNSTVSTP-ELPRKSGSTSTQMSLQAES 366
 DB 474 LGPPSGP---PDEEDGKEGST--VP---RAATYASPTKRGECTPAKPTTSPSLSSDA 525
 QY 367 KAT-----ITPSSG-----VISKFNSTSSATQAFDSSA-----VVF 401
 DB 526 PITSAPLKLMIAPSGSGVWRREPSIHHTAASGPQEPAGDSSVAITQNNQDQKLLFY 585
 QY 402 EVSTAVVAVILTMVLGLVKLCFHESPSQPKESMGPPGLESDEPPAA 451
 DB 586 ILCTVAAILLLAL-ALGL--LVYRRRAKREKK-----EKKPQNA 625

RESULT 9
 ID 000274 PRELIMINARY; PRT; 652 AA.
 AC 000274;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE C1QR(P).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-9719258; PubMed-9047234;
 RA Neomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
 RT cDNA cloning and primary structure analysis of C1QR(P), the human
 RT C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro".
 RL Immunity 6:119-129(1997).
 DR EMBL: U94333; AAB53110.1;
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR001304; Iectin_C.

DR Pfam: PF00008; EGF. 5.
 DR Pfam: PF00059; lectin_c. 1.
 DR SMART: SM00034; CLECT. 1.
 DR SMART: SM00179; EGF_CA. 3.
 DR SMART: SM00001; EGF_Like. 2.
 DR PROSITE: PS00010; ASX_HYDROXYL. 3.
 DR PROSITE: PS00041; C_Type_Lectin_2. 1.
 DR PROSITE: PS01186; EGF_2. 3.
 DR PROSITE: PS01187; EGF_CA. 3.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SO SEQUENCE 652 AA; 68577 MW; B/EBB5E5714A75 CRC64;

Query Match 11.8%; Score 307; DB 4; Length 652;
 Best Local Similarity 21.4%; Pred. No. 7.7e-17;
 Matches 139; Conservative 81; Mismatches 206; Indels 224; Gaps 28;

1 MRAFAFLCLMQLMPGGGEGHPADAGCSASGACSLHATMKROAEACILRGCA 60
 1 MATSKMLLLLLLTLPAGAGADTEAVVCGT-ACYTAHSGKLSAQAQNHQNGN 59
 61 LSTVRAGAE--LRVLA--LRAGPGGGSKDLLFWALERRSHCTLENEPLRGFSW 115
 60 LATVSKKEPAHGVVLAQOLRLREALTARMSK--FWIGLOREKGLDSEPLKGF 116
 116 LSSDGLSEDTLOWVEEPORNSCTARRCA--VLATGVEP--AGKEMRC-----HL 164
 117 VG--GGEDTPYSNMHKLKNSCKISKRCVSLDLISPLLPNRLPKMSEGGSPSGS 173
 165 RANGYLCKYQFEVLCPPARPGASNLRYAPFOLHSAALDSPGTEVSA----- 214
 174 NIEGVCKEFKGMCRPLALGPGQVITTFPOTTSSLEAVPPASANAACGGGKDET 233
 215 --LCGQLP-----ISVCTADEIGARD-----KL 238
 224 QSHFLCKEKAPDVEFDGSSGSLCVSPKYGCFNNGGCHODCFEGGSGSLCGRGR 233
 239 SGDVLT-----CPGRGY-----LRAGKCAELP 260
 294 LIDLVTCAARNPCSSPCRGATVLCPHGKNYTCRCPOGYOLDSSQLDVCVDECDSP 353
 261 --NCUDDLGFCACATGFE-----278
 354 CAQECVMPGFGFRCCEWVGEPGPGEGACQDVECALGSPCAQCGTNDGFSHCSE 413
 279 --LGRGRC--VTSGEGP-----TLGTVTRRPAT 309
 414 GYVLAGEGDCQDVECVPGGGLCDSLCFNTGSGFHCGLPGWVLAENGVSCTGMPVS 473
 310 ATSPVQRTMPPIVD--EKLGTPLVPEODNSTYVIP-ELPRWGSGSTMTLOMSLOAES 366
 474 LGPSGP--PDEEDKGEKGGT--VP--RATASPTGPECTPKATPTTSRPLSSDA 525
 367 KAT-----ITPSGS-----VSKFNSTTSATPQAFDSSA-----VVEI 401
 526 PITSAPLKLAPSGSSGVNREPSIHATASGPOEPAGDSSVAQONNDGTGOKLLIFY 585
 402 FVSTAVVAVIILMTVLGLVKLCFHESSSPRKESMGPGLSEDEPA 451
 586 ILGTVAAILLIAL-ALGL-LYKRRRAKREK-----EKKPQNA 625

RESULT 10
 ID 09J126 PRELIMINARY; PRT; 643 AA.
 AC 09J126;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
 GN C10BP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTER; TISSUE=LUNG;
 RX MEDLINE=20507883; Pubmed=10934210;
 RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.,
 RT "Molecular and Cellular Properties of the Rat A4 Antigen, a C-type
 lectin-like Receptor with Structural Homology to Thrombomodulin.";
 RL J. Biol. Chem. 275:34382-34392(2000).
 DR EMBL: AF160978; AAF0402.1;
 DR HSSP: P35353; IEMN.
 DR Interpro: IPR000152; Asx_hydroxy1.
 DR Interpro: IPR000561; EGF_Like.
 DR Interpro: IPR001881; EGF_CA.
 DR SMART: SM00034; CLECT. 1.
 DR SMART: SM00179; EGF_CA. 3.
 DR SMART: SM00001; EGF_Like. 2.
 DR PROSITE: PS00010; ASX_HYDROXYL. 3.
 DR PROSITE: PS00041; C_Type_Lectin_2. 1.
 DR PROSITE: PS01186; EGF_2. 3.
 DR PROSITE: PS01187; EGF_CA. 3.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SO SEQUENCE 643 AA; 68780 MW; FB23D742E02A4931 CRC64;

Query Match 11.7%; Score 305; DB 11; Length 643;
 Best Local Similarity 20.8%; Pred. No. 1.1e-16;
 Matches 129; Conservative 66; Mismatches 185; Indels 240; Gaps 23;

7 LCLIMQLMPGGGEGHPADAGCSASG-ACYSLHATMKROAEACILRGALSTVR 65
 11 LGLILO-LWAG-----AAADSEAVYCEGTACTAYAHWKLISAQAQNHQNGN 63
 66 AGAELRAV--LALRAPGGGGSKDLLFWALERRSHCTLENEPLRGFSWISDPG 122
 64 SEEARHVOEALQOLTKKAP-SEKTKIGFWIGLOREKGTCTYHDLPMKFSWVG--GG 119
 123 LESDTLOWVEEPORNSCTARRCAVLATGVEP--AGKEMRCNL-----RANGYL 171
 120 EDTYSNMVYKSKSSCSLRCVSLDLISLKPSPSHLPKMHESPCCGPPDAPGNSIBFLC 179
 172 KYQEVLCPPARPGASNLRYAPFOLHSAALDSPGTEVSAALC-----216
 180 KFNKGMCSPLALGPGQLITTTTFQATTSKAVPASAANVAGDDEASKTNYLCKE 239
 217 --RGQLPIS-----VTCTA 228
 240 TTAGVFWHSGSLPVCVSPKFGCSFNNNGCCQDCEGSGSFRGCGRPRLLDVLTCAS 299
 229 -----228
 300 RNPSSNPCTGGGCHSVPLSENTGHCPRGYOLDSSVYHCVYDIDCEDSPCDQECINP 359
 229 -----228
 360 GGFHCCEWYGSSGSKKEACEVDCTAAYSPCAGCGTNTDGFSCSKEGTINSGKS 419
 256 --CAELPNCI--DDL--GFCACATGPELCKGRSCVTGEGQPTLGG--TGVP 302
 420 TQCEDIDECAGNCDTLCTINTDGSFRCGCPAGFELAPNVSC-----TRGSFSELP 471
 303 TRRPATATSPVQRTMPPIR-----VDENKLGTPLVPEODNSTYVIPETPRGKS 351
 472 ARPPQKEDKGDKESVPLTEMPGSLNGSKDVSNNAQTITDLSIOSSSTASVP-----524
 352 QSTMTLSLOMSLOAESK-----ATTPSGSVTSKFNSTTSATPQAFDSSS--AVVEI 401
 525 -----LEIEVSSSEASDVLIDGLYLPPTTSGHQ--PHHEDSVPAHSDSDPDGOKLLIFY 576
 402 FVSTAVVAVIILMTVLGLV 421

Query Match 11.6%, Score 301, DB 11, Length 643;
 Best Local Similarity 20.6%, Pred. No. 2,3e-16;
 Matches 128; Conservative 67; Mismatches 185; Indels 240; Gaps 23;

```

OY 7 TLLMOLMPGPGGGEHPTADRAGCSAG-ACYSLHHTMKRQAEEACILRGSLSTVR 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11 LGLLGG-LWAG-----AAADSEAVVCECTACTYAHMKLSAAEOKHRCNENGMVLATVK 63
OY 66 AGAELRAV---LALLRAGPREGGSKDLLFWALERRSHSCTLENEPLRGFSWLSDDCG 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 SEEARHVOEALQDLTKRAP-SETKICKFMIGLORREKCKCTYHDLPMKGFSSWV- 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 123 LESDLOWVEERORCTARRCAVLQATGVGP-AGKEMRCHL-----RANGVLC 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 120 EDTTYSNNYKASRSCISCKRCVSLDLSTLRPHSHLPKMHESPCTPDADGNSIEGFLC 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 172 KYQEVILCPADPRGAASMLSYRAPFOLHSAALDESPCTEVSALC----- 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 180 KFNFKMGCSPLALGPGQLTFTTPFOATSSLKAVPFASVANVOCGDAESKTYLCKE 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 217 -----RGOLPIS-----VTCIA 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 TTAGVFHGGSGPLCVSEKFCSCFNNNGCQDCEGEGDGFRCGRGFRLLDILVTCAS 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 229 ----- 228
DB 300 RNPCCSNPCTGGGCHSVPLSENTHCHPRGYOLDSQVHVDIDECEDSCDECIINTP 359
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 229 -----DELGARD-----KLSDVLCPCPGRYLRAG-- 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 GGFHCEGWYQSSGSEACEDEVDCTAAVSPCAQCTINTDGSFYSCKEGYTMSGDS 419
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 255 -KCAELPNCI---DDL-----GGFACECATFELGKGRSCVTSGGGQPLTG--IGVP 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 TCCEDIDECIGNPCDITCINTDGSFRCGCPAGFELAPNGVSC-----TRGSMFSELP 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 303 TRRPATATSPVQRTPIR-----VDEKLEETPLVPQDMSVTSIPEIPRWGS 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 472 ARPKQKDKDGKESYVLEMPGLNSKDVSNRAQTDLISQSDSTASVP----- 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 352 QSTWSTLMSLOESK-----ATTPSGSVISKFNSTTSATPOAFDSSS---AVFT 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 525 -----LEIVSSPASVWMLDLGTYLPTTSGHSQ--PTHEDSVPAHSDSDTDQKLLFY 576
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 402 EVSTAVVAVILMTVVLGV 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
577 ILGTVAISILLAL-ALGLL 595
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

```

OY 09UC32 PRELIMINARY; PRT; 468 AA
AC 09UC32;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE THROMBOMODULIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE.
RA MEDLINE=93293792; PubMed=8390446;
RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;
RT "Urinary thrombomodulin, its isolation and characterization."
RL J. Biochem. 113:433-440(1993).
DR HSSP; P07204; 120.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.

```

```

DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00059; Lectin_C; 1.
DR PRINTS; PR00907; THROMBOMODULIN.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS50041; CTYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS01187; EGF CA; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat
SEQUENCE 468 AA; 49444 MW; 4BFEBE98EFB86A40 CRC64;

```

Query Match 8.7%, Score 227, DB 4, Length 468;
 Best Local Similarity 25.7%, Pred. No. 1.9e-10;
 Matches 85; Conservative 36; Mismatches 120; Indels 90; Gaps 18;

```

OY 16 PGEGGHPADADRAGCSAGACYSLHHTMKRQAEEACILRGALSTVRAGAEALRAVIA 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 POEGSGQCVSHD-----CEALYGPATFPLNASQICDGLRGHLMTVSSVAADVISI 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 76 LVRAGPGGGSKDLFWALE-----RRSHCTLENEPLRGFSWLSDDPGLESDT 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 56 LFN---GDGSGVGRRL-WIGQLPPEGGDPKRLG-----PLRGFWYTGDNNTSYS-- 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 128 LOWVEEROR-----CTARRCAVLQATGVGPAGKEMKCHLRANGVLCRYQEVILCP 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 103 -RWARLDLNGAPLGPCLCVASAA--EATVSEPI-WEEQCCYKADGFLCEHFPAIATR 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 -APRGAAS--NLGYRAPFOLHSAALDESPCTEVSALCRQLPISTCIAD----- 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 159 PLAVEFGAANAASVITGYGPFARAGDFQALPVGSSAAV--APGLQLMCTAPGAVYG 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 230 -----EIGARD-----KLSDVLCPCP-----GRYLRAG--KCAEL-- 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 216 HWAREAPGAMDCSEVNGGCEHACNAIPGAPROCPCPAGALQADGRCTASATQSCNDLCE 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 260 -----PNCIDDLGFACECATFELGKGRSC 286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 276 HECVFN-PPQPSYSQMCCTGYRLADQHRG 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

```

OY 035370 PRELIMINARY; PRT; 577 AA.
AC 035370;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCB1_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RL Submitted (Sep-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022743; AAB80760.1;
DR EMBL; AF022742; AAB80923.1;
DR HSSP; P07204; 1FGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Lectin_C.

```

DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF_5.
 DR Pfam: PF00059; Lectin_C; 1.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00179; EGF_CA; 2.
 DR SMART: SM00001; EGF_Like; 3.
 DR PROSITE: PS00010; ASX_HYDROXYL; 2.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 577 AA; 6184 MW; 0BE764CBF1855F CRC64;

Query Match
 Best Local Similarity 7.4%; Score 193; DB 11; Length 577;
 Matches 77; Conservative 37; Mismatches 120; Indels 70; Gaps 17;

QY 37 CYSLHATMKRQAEEACILRGALSTVRAGAEIARAVALLLRAGPGPGGSKDLLFWVAL 96
 DB 35 CFALFDPPVFELASQACRLQGLHMTVRSSVADVISLIV-----SDSSMDSRPWIGL 88

QY 97 ERRRSNC--TLENPLRGFWLSSDPGLESPTLQWVEEPRS-----CTARCAVLQ 147
 DB 89 QLPQG--CGDPVHLGLNGFGQWVGDN--HTSYSRMARPDQSPPLGCLVYSTATEA 144

QY 148 ATGVEPAGKEMKCHLRANGYLCKYQFEVLCPAPRP-----GAASNLRYAPFOL 198
 DB 145 APG--EPA--WEKPCENETGFLCEYFAFC---RLRNTDRPBEAHISSTNTPLGV 198

QY 199 HSAALDFSPGTEVA-----LCR-----GQLPISVT---CIADEIGARW--DKL 238
 DB 199 SGADFQTLPIGSSATVAPFGLIELVCRALPQTSBGHMTREVTGAWNCSEVNGCCYMCNRS 258

QY 239 SGDVLPCCP--GRTLR--GK-----CAEL-----PNCILDGSGFACGATGFEIGKD 282
 DB 259 ANGPCVCPGSGDLQADGRCACAPVAOLCNELCOHFCVNNSDVPGSYSCMCETGYOLAAD 318

QY 283 GRSC 286
 DB 319 GHRC 322

RESULT 15

Q9HCY3 PRELIMINARY; PRT; 184 AA.

AC Q9HCY3: 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)

DI 01-DEC-2001 (Tremblrel. 19, last annotation update)

DE PUTATIVE (FRAGMENT).

GN PREDI2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eultheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,

RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,

RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,

RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,

RA Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,

RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,

RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,

RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,

RA Schafte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,

RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,

RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,

RA Lehmach H., Reinhardt R., Vaspou M.L.,

RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RA EMBL: AL163217; CAB90388.1; -

DR HSSP: P06734; 1HLI.
 DR InterPro: IPR001304; lectin_C.
 DR Pfam: PF00059; lectin_C; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 FT NON_TER 1 1
 FT NON_TER 184 184
 SQ SEQUENCE 184 AA; 20819 MW; DD6B28E6A91BD43 CRC64;

Query Match
 Best Local Similarity 5.5%; Score 143.5; DB 4; Length 184;
 Matches 48; Conservative 21; Mismatches 56; Indels 47; Gaps 9;

QY 37 CYSL--HATMKR--QAEEACILRGALSTVRAGAEIARAVALLLRAGPGPGGSKDLL 91
 DB 12 CYKMYFHELSRVSFQEARLACESEGVLLSTLENAEQKILIESMLQNTKPKGISDGD 71

QY 92 FWALERRRSHCTLENPLRGFWLSSD--PGLESPTLQW-----VEEPRS 137
 DB 72 FWIGL-----WRNGDQGTSGACPDLYQWSDGNSQYRMWYDDEP--S 111

QY 138 CTARCAVL--QAT-----GVEPAGKEMKCHLRANGYLCKYQFEVLCPAP 182
 DB 112 CGSEKCVMYHQPANPGLGPGVLYQWNNDRCNMKNH-YICKYPEINPTAP 162

Search completed: August 13, 2002, 09:37:54
 Job time: 238 sec

Tue Aug 13 09:56:31 2002

us-09-902-713b-96.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      August 13, 2002, 09:34:26 ; Search time 34.18 seconds
              (without alignments)
              1592.339 Million cell updates/sec
```

Title: US-09-902-713B-96

Sequence: 1 MRPAFALCLLWQALWPGG.....LRDRAEGALLAESPLGSSDA 49C

Scoring table: OLIGO

Searched: 747574 seqs, 111073796 residues

0 size : 0

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 150 summaries

Database : A_Geneseq_032802:*

| | |
|-----|--|
| 1: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT * |
| 2: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT * |
| 3: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT * |
| 4: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT * |
| 5: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT * |
| 6: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT * |
| 7: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT * |
| 8: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT * |
| 9: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT * |
| 10: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT * |
| 11: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT * |
| 12: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT * |
| 13: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT * |
| 14: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT * |
| 15: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT * |
| 16: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT * |
| 17: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT * |
| 18: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT * |
| 19: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT * |
| 20: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT * |
| 21: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT * |
| 22: | /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT * |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query % | Match | length | DB | ID | Description |
|--------|-----|-------|---------|-------|----------|----|----|--------------------|
| 1 | 490 | 100.0 | 490 | 20 | AAV13360 | | | Amino acid sequenc |
| 2 | 490 | 100.0 | 490 | 21 | AAB33460 | | | Human PRO269 prote |
| 3 | 490 | 100.0 | 490 | 21 | AAB24403 | | | Human PRO269 prote |
| 4 | 490 | 100.0 | 490 | 21 | AAV95016 | | | Human secreted pro |
| 5 | 490 | 100.0 | 490 | 22 | AAU12349 | | | Human PRO269 poly |
| 6 | 490 | 100.0 | 490 | 22 | AAE06596 | | | Human protein hav |
| 7 | 296 | 60.4 | 331 | 21 | AAB41847 | | | Human ORFX ORF161 |
| 8 | 274 | 55.9 | 344 | 21 | AAV91410 | | | Human secreted pro |
| 9 | 242 | 49.4 | 428 | 22 | AAB80228 | | | Human PRO269 prote |
| 10 | 139 | 28.4 | 175 | 21 | AAV95031 | | | Human clone vp15.1 |
| 11 | 84 | 17.1 | 189 | 21 | AAV91349 | | | Human secreted pro |

[illegible]

| | | | | | | |
|-----|---|-----|-----|----|----------|----------------------|
| 85 | 7 | 1.4 | 144 | 22 | AAU56983 | Propionibacterium |
| 86 | 7 | 1.4 | 147 | 22 | AAU41537 | Propionibacterium |
| 87 | 7 | 1.4 | 147 | 22 | AAU87796 | Human immune/haema |
| 88 | 7 | 1.4 | 148 | 22 | AAU01882 | Human polypeptide |
| 89 | 7 | 1.4 | 153 | 21 | AAU65656 | Propionibacterium |
| 90 | 7 | 1.4 | 159 | 21 | AAU14119 | Human secreted pro |
| 91 | 7 | 1.4 | 159 | 22 | AAU85241 | Human secreted pro |
| 92 | 7 | 1.4 | 177 | 22 | AAU09342 | Novel human diagno |
| 93 | 7 | 1.4 | 218 | 12 | AAU11000 | Antigenic peptide |
| 94 | 7 | 1.4 | 229 | 21 | AAU41733 | Human ORFX ORF1497 |
| 95 | 7 | 1.4 | 238 | 10 | AAU93696 | Antigenic polypept |
| 96 | 7 | 1.4 | 244 | 22 | AAU32672 | Novel human secret |
| 97 | 7 | 1.4 | 246 | 22 | AAU58306 | Propionibacterium |
| 98 | 7 | 1.4 | 267 | 21 | AAU52135 | Arabidopsis thalia |
| 99 | 7 | 1.4 | 288 | 21 | AAU27461 | Arabidopsis thalia |
| 100 | 7 | 1.4 | 290 | 22 | ABG12645 | Novel human diagno |
| 101 | 7 | 1.4 | 291 | 22 | ABG20856 | Human AFP protein |
| 102 | 7 | 1.4 | 303 | 22 | AAU81311 | Human ORFX ORF3020 |
| 103 | 7 | 1.4 | 312 | 21 | AAU43256 | Drosophila melanog |
| 104 | 7 | 1.4 | 319 | 17 | AAU00628 | Human Homeobox gen |
| 105 | 7 | 1.4 | 320 | 22 | ABU61055 | Human olfactory re |
| 106 | 7 | 1.4 | 330 | 20 | AAU66020 | Murine homeobox HO |
| 107 | 7 | 1.4 | 332 | 22 | AAU72076 | Novel human diagno |
| 108 | 7 | 1.4 | 333 | 20 | AAU66019 | TANGO 300 signal-1 |
| 109 | 7 | 1.4 | 333 | 22 | ABU19326 | Human diagnostic p |
| 110 | 7 | 1.4 | 336 | 20 | AAU66365 | Human DNX toll-11 |
| 111 | 7 | 1.4 | 341 | 22 | AAU61462 | Human diagnostic p |
| 112 | 7 | 1.4 | 346 | 22 | AAU19442 | Human cancer assoc |
| 113 | 7 | 1.4 | 353 | 21 | AAU43736 | Human colon cancer |
| 114 | 7 | 1.4 | 353 | 22 | AAU37374 | Drosophila melanog |
| 115 | 7 | 1.4 | 359 | 22 | ABU66298 | Human secreted pro |
| 116 | 7 | 1.4 | 361 | 21 | AAU23618 | Human AFP protein |
| 117 | 7 | 1.4 | 361 | 22 | AAU61399 | Human AFP protein |
| 118 | 7 | 1.4 | 361 | 22 | AAU61421 | Human TANGO 300 pr |
| 119 | 7 | 1.4 | 384 | 22 | AAU99127 | Human G protein-co |
| 120 | 7 | 1.4 | 384 | 22 | AAU66727 | Human 16405 protei |
| 121 | 7 | 1.4 | 386 | 22 | AAU79926 | Corynebacterium gl |
| 122 | 7 | 1.4 | 397 | 20 | AAU35476 | Chlamydia pneumonia |
| 123 | 7 | 1.4 | 399 | 22 | AAU71595 | Arabidopsis thalia |
| 124 | 7 | 1.4 | 402 | 21 | AAU54363 | Arabidopsis thalia |
| 125 | 7 | 1.4 | 402 | 22 | AAU20527 | Human secreted pro |
| 126 | 7 | 1.4 | 406 | 22 | ABU11445 | Novel human diagno |
| 127 | 7 | 1.4 | 408 | 20 | AAU41111 | Human TANGO 129 (T |
| 128 | 7 | 1.4 | 414 | 22 | AAU32999 | Novel human secret |
| 129 | 7 | 1.4 | 418 | 22 | AAU91830 | C glutamicum prote |
| 130 | 7 | 1.4 | 429 | 22 | ABU09437 | Novel human diagno |
| 131 | 7 | 1.4 | 429 | 22 | AAU71961 | Human TGF-beta rec |
| 132 | 7 | 1.4 | 430 | 20 | AAU41110 | Human TGF-beta rec |
| 133 | 7 | 1.4 | 430 | 21 | AAU70785 | Human tumour necro |
| 134 | 7 | 1.4 | 430 | 22 | AAU09901 | Human tumour necro |
| 135 | 7 | 1.4 | 430 | 22 | AAU05517 | Human protein sequ |
| 136 | 7 | 1.4 | 430 | 22 | AAU95627 | Human polypeptide, |
| 137 | 7 | 1.4 | 434 | 22 | AAU93638 | Human PRO polypept |
| 138 | 7 | 1.4 | 440 | 22 | AAU29216 | Human PRO polypept |
| 139 | 7 | 1.4 | 440 | 22 | AAU87592 | Human PRO polypept |
| 140 | 7 | 1.4 | 460 | 22 | AAU98288 | Escherichia coli p |
| 141 | 7 | 1.4 | 461 | 22 | AAU62188 | C. pneumoniae coll p |
| 142 | 7 | 1.4 | 466 | 20 | AAU50341 | Human K6.2 protei |
| 143 | 7 | 1.4 | 466 | 22 | ABU11444 | Novel human diagno |
| 144 | 7 | 1.4 | 481 | 19 | AAU83216 | Human h-RAB-R... HO |
| 145 | 7 | 1.4 | 481 | 22 | AAU50342 | Murine K6.2 prote |
| 146 | 7 | 1.4 | 481 | 22 | AAU04758 | Human vesicle traf |
| 147 | 7 | 1.4 | 502 | 22 | ABU69102 | Drosophila melanog |
| 148 | 7 | 1.4 | 510 | 22 | AAU84250 | Amino acid sequenc |
| 149 | 7 | 1.4 | 522 | 22 | AAU84250 | R. marinus bacteri |
| 150 | 7 | 1.4 | 526 | 21 | AAU54362 | Arabidopsis thalia |

ALIGNMENTS

| | |
|----|--|
| ID | AAU13360 standard; Protein: 490 AA. |
| XX | AAU13360; |
| AC | 25-JUN-1999 (first entry) |
| XX | Amino acid sequence of protein PRO269. |
| DE | Secreted protein; transmembrane protein; human; enterocolitis; |
| XX | Zollinger-Ellison syndrome; gastrointestinal ulceration; |
| KW | congenital microvillus atrophy; skin disease; cell growth; |
| KW | abnormal keratinocyte differentiation; psoriasis; epithelial cancer; |
| KW | Parkinson's disease; Alzheimer's disease; ALS; neuropathy; |
| KW | fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata; |
| KW | anti-thrombotic; wound healing; tissue repair. |
| OS | Homo sapiens. |
| XX | WO914328-A2. |
| PN | 25-MAR-1999. |
| PD | 16-SEP-1998; 98MO-US19330. |
| XX | 25-NOV-1997; 97US-0066840. |
| PF | 17-SEP-1997; 97US-0059113. |
| XX | 17-SEP-1997; 97US-0059115. |
| PR | 17-SEP-1997; 97US-0059117. |
| PR | 17-SEP-1997; 97US-0059119. |
| PR | 17-SEP-1997; 97US-0059121. |
| PR | 17-SEP-1997; 97US-0059122. |
| PR | 18-SEP-1997; 97US-0059184. |
| PR | 18-SEP-1997; 97US-0059263. |
| PR | 15-OCT-1997; 97US-0059266. |
| PR | 15-OCT-1997; 97US-0061215. |
| PR | 17-OCT-1997; 97US-0062285. |
| PR | 17-OCT-1997; 97US-0062287. |
| PR | 21-OCT-1997; 97US-0062287. |
| PR | 21-OCT-1997; 97US-0062814. |
| PR | 24-OCT-1997; 97US-0062816. |
| PR | 24-OCT-1997; 97US-0063045. |
| PR | 24-OCT-1997; 97US-0063120. |
| PR | 24-OCT-1997; 97US-0063121. |
| PR | 24-OCT-1997; 97US-0063129. |
| PR | 27-OCT-1997; 97US-0063327. |
| PR | 27-OCT-1997; 97US-0063329. |
| PR | 28-OCT-1997; 97US-0063341. |
| PR | 28-OCT-1997; 97US-0063342. |
| PR | 28-OCT-1997; 97US-0063344. |
| PR | 28-OCT-1997; 97US-0063349. |
| PR | 28-OCT-1997; 97US-0063350. |
| PR | 28-OCT-1997; 97US-0063354. |
| PR | 29-OCT-1997; 97US-0063355. |
| PR | 29-OCT-1997; 97US-0063364. |
| PR | 29-OCT-1997; 97US-0063365. |
| PR | 29-OCT-1997; 97US-0063370. |
| PR | 29-OCT-1997; 97US-0063372. |
| PR | 29-OCT-1997; 97US-0063373. |
| PR | 29-OCT-1997; 97US-0063374. |
| PR | 29-OCT-1997; 97US-0063375. |
| PR | 31-OCT-1997; 97US-0063375. |
| PR | 31-OCT-1997; 97US-0063375. |
| PR | 31-OCT-1997; 97US-0063375. |
| PR | 03-NOV-1997; 97US-0064103. |
| PR | 07-NOV-1997; 97US-0064248. |
| PR | 12-NOV-1997; 97US-0064809. |
| PR | 17-NOV-1997; 97US-0065186. |
| PR | 17-NOV-1997; 97US-0065846. |
| PR | 18-NOV-1997; 97US-0065693. |
| PR | 21-NOV-1997; 97US-0066120. |
| PR | 21-NOV-1997; 97US-0066364. |
| PR | 24-NOV-1997; 97US-0066772. |
| PR | 24-NOV-1997; 97US-0066466. |
| PR | 24-NOV-1997; 97US-0066770. |
| PR | 24-NOV-1997; 97US-0066511. |


```

PR 24-NOV-1997; 97US-0066453.
XX (GETH ) GENENTECH INC.
XX Chen J, Goddard A, Gurney AL, Pennica D, Wood WL, Yuan J;
XX MPI: 1999-229533/19.
XX N-PSDB; AAX52231.
XX
XX New isolated human genes and polypeptides used in, e.g. treatment of
XX gastrointestinal ulceration
XX
XX Claim 12, Fig 36, 320pp; English.
XX
XX AAV13344-403 represent secreted and transmembrane human proteins.
XX The cDNA sequences are obtained from cDNA libraries, prepared from
XX fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
XX The encoded polypeptides have specific uses based on their homology to
XX known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
XX associated with the preservation and maintenance of gastrointestinal
XX mucosa and the repair of acute and chronic mucosal lesions
XX (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
XX ulceration and congenital microvillus atrophy), skin diseases associated
XX with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
XX cancers such as lung squamous cell carcinoma of the vulva and gliomas),
XX potent effects on cell growth and development, diseases related to
XX growth or survival of nerve cells including Parkinson's disease,
XX Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
XX for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
XX as a target for anti-tumor drugs. PRO533 may be used in the treatment
XX of Usher Syndrome or Atrophia areata. PRO269 can be used as an
XX anti-thrombotic agent; PRO287 polypeptides and portions may have
XX therapeutic applications in wound healing and tissue repair; PRO317 can
XX be used for treating problems in the kidney, uterus, endometrium, blood
XX vessels, or related tissue, e.g. in the heart of genital tract.
XX
XX Sequence 490 AA:
XX
XX Query Match 100.0%; Score 490; DB 20; Length 490;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 421 vlcifhespsppikesmgppglesdpepaalgssahctmgvkvgdcdlrdraegall 480
OY 481 AESPLGSSDA 490
DB 481 aesplgssda 490
XX
XX RESULT 2
XX AAB33460
XX ID AAB33460 standard; Protein; 490 AA.
XX
XX AAB33460;
XX
XX 29-JAN-2001 (first entry)
XX
XX Human PRO269 protein UNQ236 SEQ ID NO:233.
XX
XX Human; immune related disease; diagnosis; antiinflammatory; cardiatic;
XX dermatological; anrlarthritis; anrlrheumatic; immunosuppressive;
XX haemostatic; antihypoid; antidiabetic; nootropic; neuroprotective;
XX antianemic; hepatotropic; virucide; antipsoriatic; antiallergic;
XX antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
XX osteoarthritis; spondylarthropathy; systemic sclerosis; sarcoidosis;
XX idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
XX systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
XX autoimmune thrombocytopenia; immune-mediated renal disease;
XX demyelinating disease; hepatobiliary disease; Whipple's disease;
XX inflammatory bowel disease; gluten-sensitive enteropathy;
XX autoimmune disease; immune-mediated skin disease; allergic disease;
XX immunological disease; transplantation associated disease;
XX graft rejection; graft-versus-host-disease.
XX
XX Homo sapiens.
XX
XX WO200053758-A2.
XX
XX 14-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05841.
XX
XX 08-MAR-1999; 99WO-US05028.
XX 10-MAR-1999; 99US-0123618.
XX 12-MAR-1999; 99US-0123957.
XX 23-MAR-1999; 99US-0125775.
XX 12-APR-1999; 99US-0128849.
XX 20-APR-1999; 99WO-US08615.
XX 28-APR-1999; 99US-0131445.
XX 04-MAY-1999; 99US-0132371.
XX 14-MAY-1999; 99US-0134287.
XX 02-JUN-1999; 99WO-US12252.
XX 23-JUN-1999; 99US-0141037.
XX 20-JUL-1999; 99US-0144758.
XX 26-JUL-1999; 99US-0145698.
XX 28-JUL-1999; 99US-0146222.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US20944.
XX 15-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 05-OCT-1999; 99WO-US23089.
XX 29-OCT-1999; 99US-0162506.
XX 29-NOV-1999; 99WO-US28214.
XX 30-NOV-1999; 99WO-US28313.
XX 30-NOV-1999; 99WO-US28409.
XX 01-DEC-1999; 99WO-US28301.
XX 01-DEC-1999; 99WO-US28634.
XX 02-DEC-1999; 99WO-US28551.
XX 02-DEC-1999; 99WO-US28564.
XX 02-DEC-1999; 99WO-US28565.
XX 16-DEC-1999; 99WO-US30095.
XX 20-DEC-1999; 99WO-US30999.
XX 30-DEC-1999; 99WO-US31274.
XX 05-JAN-2000; 2000WO-US00219.

```

PR 06-JAN-2000; 2000MO-US00277
 PR 06-JAN-2000; 2000MO-US00376
 PR 11-FEB-2000; 2000MO-US03565
 PR 18-FEB-2000; 2000MO-US04341
 PR 18-FEB-2000; 2000MO-US04342
 PR 22-FEB-2000; 2000MO-US04414

XX (GETH) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W,
 PI Kabakoff RC, Lu Y, Pan Y, Pennica D, Shelton DJ, Smith V,
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M,
 DR WPI; 2000-572271/53.
 DR N-PSDB; AAC58625.

PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus.

XX Claim 33; Fig 94; 309pp; English.

CC The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central
 CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
 CC autoimmune or immune-mediated skin diseases, allergic diseases,
 CC diseases including graft rejection and graft-versus-host disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAC33414 to AAC33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.

XX Sequence 490 AA;

Query Match 100.0%; Score 490; DB 21; Length 490;
 Best Local Similarity 100.0%; P-vec No. 0;

Mismatches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRPAFLCLIMQALWPGSGGHEPTDRACSGACYSLHATMKROAEACILRGGA 60
 Db 1 mrpafalclimqalwpgsggheptdracsgacyslhatmkrgaeeccilrgga 60
 Oy 61 LSTVRACAEELRAVLALIRAGPGGSKDLFWALERRRSICTLENEPLRGFSWISDP 120
 Db 61 lstrvrageelravalliragpggskdlfwalerrrrsictleneplrgfswissdp 120
 Oy 121 GGLSEDTLQWVEPORSCTARRCAVLQATGVEPAGKEMKRCRLRANGVLGYOEVLCP 180
 Db 121 gglsestdlqwveeporsctarrcaavlgatgvepawkemrchrlrangylckygvevlcp 180
 Oy 181 APPGASNSRYAPROLISALIDSPFTEVSALCRGOLPISVTCIAETIGARDKLSG 240
 Db 181 appgasaansryaprolisaldspftevsalcrqolpistvciaetigarwdklsq 240
 Oy 241 DVLCPCGRYLRAKCAELPNCILDGFGACATGFEELGKRGRCVNSGEOPPLGTG 300
 Db 241 dvlcpcgrylrakcaelpncildgfgacatgfeelgkrgrcvnsgeopplgtg 300
 Oy 301 VPLTRPATATSPVQRTWPIVDEKLGFTPLVPEODMSVTSIPEIPWGSOSTWSTLQ 360
 Db 301 vpltrpatatvspvqrtwpiivdeklgftplvpeodmsvtsipeipwgsostwstlqm 360

Oy 361 SLQAEKATIPSGSVISKENSTTSATPOAFDSSAVFIEVSTAVVLVILTMVVLGL 420
 Db 361 slqaeakatlpsgsviskinsttsatpafidssavvifvstlavvvlviltmvlgl 420
 Oy 421 VKLCFHESPSQPRKESGPGLESDPEPALGSSAHCTNNCKVKGDCOLRRRACALL 480
 Db 421 vkcfhespsqprkessgpglesdpepalgssahctnnckvkgdcollrrracall 480
 Oy 481 AESPLGSSDA 490
 Db 481 aesplgssda 490

RESULT 3
 AAB24403
 ID AAB24403 standard; Protein: 490 AA.
 AC AAB24403;

DT 07-NOV-2000 (first entry)

DE Human PRO269 protein sequence SEQ ID NO:77.

KW Human: PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
 KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
 KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;
 KW cytosolic; gene therapy; vaccine.

OS Homo sapiens;

PN WO200032221-A2.

XX 08-JUN-2000.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1998; 98WO-US25108.

XX 16-DEC-1998; 98US-0112850.

XX 12-JAN-1999; 99US-0115554.

XX 08-MAR-1999; 99WO-US05028.

XX 12-MAR-1999; 99US-0123957.

XX 28-APR-1999; 99US-0131445.

XX 14-MAY-1999; 99US-0134287.

XX 02-JUN-1999; 99WO-US12252.

XX 23-JUN-1999; 99US-0141037.

XX 20-JUL-1999; 99US-0144758.

XX 26-JUL-1999; 99US-0145698.

XX 01-SEP-1999; 99WO-US20111.

XX 08-SEP-1999; 99WO-US20594.

XX 13-SEP-1999; 99WO-US20944.

XX 15-SEP-1999; 99WO-US21090.

XX 05-SEP-1999; 99WO-US21547.

XX 29-OCT-1999; 99WO-US23089.

XX 29-OCT-1999; 99US-0162506.

PA (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A,
 PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V,
 PI Watanabe CK, Williams PM, Wood WI;

DR WPI; 2000-412154/35.
 DR N-PSDB; AAA77570.

PT Nucleic acids encoding PRO polypeptides useful for preventing,
 PT diagnosing and treating disorders in mammals
 PT angiogenic disorders in mammals

PS Claim 72; Fig 32; 315pp; English.

CC The present invention describes nucleic acids encoding PRO polypeptides
 CC useful for preventing, diagnosing and treating disorders in mammals by
 CC cardiovascular, endothelial or angiogenic disorder in mammals by

modulating cell proliferation, angiogenesis and cardiovascularisation, and for identifying agonists and antagonists of these processes. The CC nucleic acids and the proteins they encode may be used in the CC prevention, treatment and diagnosis of diseases associated with CC. Inappropriate PRO expression such as cardiovascular, endothelial or angiogenic disorders in mammals (e.g. atherosclerosis, cancers and cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors containing them and the PRO polypeptide may be used to treat disorders associated with decreased PRO expression. AA77510 to AA77721 and CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in the exemplification of the present invention.

Sequence 490 AA:

Query Match 100.0%; Score 490; DB 21; Length 490;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPAFLCLLMQALMPGCGGHEPTADRAAGCSAGACYSLHNTMKROAAEEACILRGGA 60
D 1 mrpafalclllwgalwpgpggghptadragcsagacyslhntmrqaaeeacillrgga 60
QY 61 LSTVRAGAEIRAVLALLRAGPGGSGKDLLFWALERRSHCTLENEPLRGFSWLSDDP 120
D 61 lstrvragaelravalllragspgpggskdllfwalerrshctlenepirgfswlssdp 120
QY 121 GLESTDTLQWVEEPPORSCSTARCAVLAQATGVEPAGKEMKCHLRANGYLCKYQFEVLCP 180
D 121 glesdtllqweepqscstarccavlaqatgvepagkewmchllrangylckyqfevlcp 180
QY 181 APRGAASNLSTRAPQLHSAALDPSPTCEVSALCRGOLPISVTCIADIGARMKLSG 240
D 181 aprgaasnlsyrapqlhsaaldisptcevsalcrqlplsvtciaaigarmkls 240
QY 241 DVLCPGPGYRLAAGKCAEPLNCIDLDGFGFACATGFEIGKORSCVTSGEQPTIGGG 300
D 241 dvlcpcpgyrlagkcaelpncldldgfgfacaatgfeigkgrscvtsgeqptlggg 300
QY 301 VPTRRPATATSPVQRTWPIRVERKLGETPLVPEODNSVTISPEIPRMGSOSTMSTLQ 360
D 301 vptrrpatatspvqrtwpirvereklgetplvpeodnsvtispeiprmgsostmstlq 360
QY 361 SIQAEKATITSGSVYSKFNSTTSATQAPDESSAVVIFPSTAVVILITMVLGL 420
D 361 siqaekatitpsgsvyskfnsttsatqapdeassaavvifpstavvilltmvlgl 420
QY 421 VLICFHESSQPRKESMPGGLSEDPPEALGSSSAHCTNNGVKYGDDCLDRACGALL 480
D 421 vlcfhespsqprkesmppyglsedppeaalgssahctnnngvkvgyddclldraagall 480
QY 481 AESPLGSSDA 490
D 481 aesplgssda 490

RESULT 4

AA95016
ID AA95016 standard; Protein; 490 AA.

AC AA95016;

DT 19-JUN-2000 (first entry)

DE Human secreted protein vp15_1, SEQ ID NO:72.

KX Human; secreted protein; cancer; tumour; cardiovascular disorder;

KM blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;

KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;

OS neurodegenerative disease; asthma; contraceptive.

XX Homo sapiens.

PN WO200011015-A1.
XX 02-MAR-2000.
XX 24-AUG-1999; 99WO-US19351.
XX 24-AUG-1998; 98US-0097638.
XX 24-AUG-1998; 98US-0097659.
XX 09-SEP-1998; 98US-0099618.
XX 28-SEP-1998; 98US-0102092.
XX 25-NOV-1998; 98US-0109978.
XX 23-DEC-1998; 98US-0113645.
XX 23-DEC-1998; 98US-0113646.
XX 23-AUG-1999; 99US-0379246.

(ALPH-) ALPHAGENE INC.

PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

DR WPI; 2000-224657/19.

PT New secreted or transmembrane proteins and polynucleotides encoding
PT them, useful for treating neurodegenerative disorders, autoimmune
PT diseases and cancer -

PS Claim 81; Page 331-332; 357pp; English.

The invention relates to 40 human secreted proteins (AA94981-995020),
and CDNA sequences encoding them (AA23423-423462). The secreted
proteins of the invention include those that are thought to be only
partially secreted, i.e., transmembrane proteins. The proteins of the
invention may exhibit one or more activities selected from the following:
CC cytokine activity; cell proliferation; differentiation; immune
CC modulation; haematopoiesis regulation; tissue growth activity;
CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
CC and thrombolytic activity; anti-inflammatory activity; and tumour
CC inhibition activity. The proteins may be administered to patients as
vaccines, and the nucleotides may be used as part of a gene therapy
CC regime. Diseases or conditions that may be treated using the proteins or
CC nucleotides of the invention include autoimmune diseases; genetic
CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
CC fungal and viral infections, especially HIV; multiple sclerosis;
CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
CC insulin dependent diabetes mellitus; and allergic reactions such as
CC asthma and anaemia. They may also be used for treating wounds, burns,
CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
CC disease, Parkinson's disease, Huntington's disease and amyotrophic
CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
CC additionally be useful as contraceptives. Nucleic acid sequences of the
CC invention may be used in chromosome mapping, and as a source of
CC diagnostic primers and probes. The present sequence represents one of the
CC 40 proteins of the invention.

Sequence 490 AA:

Query Match 100.0%; Score 490; DB 21; Length 490;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPAFLCLLMQALMPGCGGHEPTADRAAGCSAGACYSLHNTMKROAAEEACILRGGA 60
D 1 mrpafalclllwgalwpgpggghptadragcsagacyslhntmrqaaeeacillrgga 60
QY 61 LSTVRAGAEIRAVLALLRAGPGGSGKDLLFWALERRSHCTLENEPLRGFSWLSDDP 120
D 61 lstrvragaelravalllragspgpggskdllfwalerrshctlenepirgfswlssdp 120
QY 121 GLESTDTLQWVEEPPORSCSTARCAVLAQATGVEPAGKEMKCHLRANGYLCKYQFEVLCP 180
D 121 glesdtllqweepqscstarccavlaqatgvepagkewmchllrangylckyqfevlcp 180
QY 181 APRGAASNLSTRAPQLHSAALDPSPTCEVSALCRGOLPISVTCIADIGARMKLSG 240

```

DB 181 aprrgaasnlstypafqqlhsaalfdfpptevalerqgqilpsvcladeigardklsq 240
OY 241 DVLCPGGRYLRAGKCAELPNCIDLGAFACATGFEELGDKGRSCVTSGEQPTLGIGTG 300
DB 241 dvlcpgrylragkcaelpncidlgafacatgfeelgdkgrscvtsgeqptlgigtg 300
OY 301 VPTRRPATATSPVPORTWIRVDEKLGEPPLVPEDNSVTSIPEIRPWGSOSTSTLMOM 360
DB 301 vptrrpatatspvportwirvdeklegpplvpednsvtsipeirpwgsoststlmom 360
OY 361 SIQAESKATITPSGSVISFNSTSSAPPQAFDSSSAVVFVSTAVVVLITMTVLGL 420
DB 361 siqaeskatitpsgsvisfnstssappqafdsassaavvfsvstavvvlitmtvlgl 420
OY 421 VKLCFHESSPSOPRKESMGPPLESDPPALGSSSACTNNGKVGCCDLRDRAEGALL 480
DB 421 vkcfhespsoprkessmgpplesdppalgsassaactnngkvgcdldrdraegall 480
DB 481 AESPLGSSDA 490
DB 481 aesplgssda 490
RESULT 5
AAU12349
ID AAU12349: standard; Protein: 490 AA.
AC AAU12349:
XX 24-OCT-2001 (first entry)
DE Human PRO269 polypeptide sequence.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy.
XX
OS Homo sapiens.
XX
PN HQ200140466-A2
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000OMO-US32678.
XX
F 01-DEC-1999; 99WO-US28301.
F 02-DEC-1999; 99WO-US28634.
F 02-DEC-1999; 99WO-US28551.
F 02-DEC-1999; 99WO-US28564.
F 02-DEC-1999; 99WO-US28565.
F 09-DEC-1999; 99WO-US170262.
F 16-DEC-1999; 99WO-US30095.
F 20-DEC-1999; 99WO-US30911.
F 20-DEC-1999; 99WO-US30999.
F 30-DEC-1999; 99WO-US31243.
F 06-JAN-2000; 2000OMO-US00277.
F 06-JAN-2000; 2000OMO-US00376.
F 11-FEB-2000; 2000OMO-US03565.
F 18-FEB-2000; 2000OMO-US04341.
F 18-FEB-2000; 2000OMO-US04342.
F 22-FEB-2000; 2000OMO-US04414.
F 24-FEB-2000; 2000OMO-US04914.
F 24-FEB-2000; 2000OMO-US05004.
F 01-MAR-2000; 2000OMO-US05601.
F 21-MAR-2000; 2000OMO-US07377.
F 30-MAR-2000; 2000OMO-US07532.
F 17-MAY-2000; 2000OMO-US08439.
F 17-MAY-2000; 2000OMO-US13705.
F 30-MAY-2000; 2000OMO-US14042.
F 02-JUN-2000; 2000OMO-US15264.

```

```

PR 10-NOV-2000; 2000OMO-US30873.
PA (GETH) GENENTECH INC.
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Flivaroff E, Gao W,
PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
PI Smith V, Stewart JA, Tumas D, Watanabe CK, Wood WL, Zhang Z,
DR WPI: 2001-408281/43.
DR N-PSDB; AAS21421.
XX
PT Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical
XX
PS Claim 12; Fig 356; 813pp; English.
XX

```

```

CC AAU12172-AAU12446 represent novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
SQ Sequence 490 AA;

```

```

Query Match 100.0%; Score 490; DB 22; Length 490;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MRPAALCLLWQALMPGGGGEHPADBAGCSACACYSLHHATMKROAEACILNGGA 60
DB 1 mrrpaalcllwqalmpggggehpadbagsacacsylhhatmkroaeacilngga 60
OY 61 LSTVRAEELRAVALLRAGGPGGSKDLFWALERRSHCHLENEPIRGSMSSDP 120
DB 61 lstvraeelravallragpgpggskdlfwalerrshchleneprlgsmssdp 120
OY 121 GGLSDTLQWVEEPQNSCTARCAVLQATGVEPAGKREKCHLRANGYLCKYQFEVLC 180
DB 121 gglstdtlqwveepqnsctarcavllqatgvepagkrekchlrangylckyqfevllc 180
OY 181 APRGAASNLSTYPFQQLHSAALDFSPPTREVALRCGQLPISTYCTADETGARMKLSG 240
DB 181 aprgaasnlstypfqlhsaaldfspptrevalrcgqlpistytctadegtarmkls 240
OY 241 DVLCPGGRYLRAGKCAELPNCIDLGAFACATGFEELGDKGRSCVTSGEQPTLGIGTG 300
DB 241 dvlcpgrylragkcaelpncidlgafacatgfeelgdkgrscvtsgeqptlgigtg 300
OY 301 VPTRRPATATSPVPORTWIRVDEKLGEPPLVPEDNSVTSIPEIRPWGSOSTSTLMOM 360
DB 301 vptrrpatatspvportwirvdeklegpplvpednsvtsipeirpwgsoststlmom 360
OY 361 SIQAESKATITPSGSVISFNSTSSAPPQAFDSSSAVVFVSTAVVVLITMTVLGL 420
DB 361 siqaeskatitpsgsvisfnstssappqafdsassaavvfsvstavvvlitmtvlgl 420

```

OY 421 VKICHPSSOPRKESMPGLESDPEPALGSSSAHCTNNGVGVGCDLDRRAEGALL 480
 DB 421 VKICHPSSOPRKESMPGLESDPEPALGSSSAHCTNNGVGVGCDLDRRAEGALL 480
 OY 481 AESPLGSSDA 490
 DB 481 aesplgssda 490
 RESULT 6
 AAE06596
 ID AAE06596 standard; Protein: 490 AA.
 AC AAE06596;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human protein having hydrophobic domain, HP03983.
 XX
 KW human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contraceptive; antiinfectivity; antiinflammatory.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= Signal_peptide
 FT Protein 23..490
 FT /note= "Mature human protein with hydrophobic domain"
 XX
 PN MO200149728-A2.
 PD 12-JUL-2001.
 PF 28-DEC-2000; 2000WO-JP09359.
 PR 06-JAN-2000; 2000JP-0000585.
 PR 06-JAN-2000; 2000JP-0000588.
 PR 11-JAN-2000; 2000JP-0002299.
 PR 03-FEB-2000; 2000JP-0026862.
 PR 03-MAR-2000; 2000JP-0058367.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI: 2001-418355/44.
 DR N-PSDB: AAD12591.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation
 XX
 PS Claim 1; Page 110-111; 563pp; English.
 XX
 CC The present sequence is human protein with hydrophobic domain,
 CC HP03983. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC polypeptides may also be used as antigens in the production of antibodies
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional

CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.
 XX
 SO Sequence 490 AA:
 Query Match 100.0%; Score 490; DB 22; Length 490;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MRPAFLCLLMQALMPGCGGHEPTADRAGCSAGACYSLHHTMKROAEACILRGCA 60
 DB 1 mpaflclllmqalwpgpggshptadragcsagacyslhhtmkrgaaeacillrga 60
 OY 61 ISTVAGAEIRAVLALLRAGPGGGSKDLFWVALERRSHCTLENEPLRGFSMLSDP 120
 DB 61 istvagaeltavalllragpgpggskdlfwvalerrshctleneplrgfswlssdp 120
 OY 121 GLESDTLQWVEEPORSCSTARCAVLQATGVPEPGWKEMRCHLRANGYLCKYQFEVLCP 180
 DB 121 glesdttlqwveeporscstarccavllqatgvpepgwkemrchlranngylckyqfevlcp 180
 OY 181 APRPAASNLSTRAAPQLHSAALDFSPCTEVSALCRGLPISTVCIADETCARMDKLSG 240
 DB 181 aprpaasnlsyrapqlhsaaldfspctevsalcrglplstvciaadelgarvdklsg 240
 OY 241 DVLCPGGRYLKAGCAELPNCIDLDGFCACAGFEELGKRGSCVSGEGOPTLGNG 300
 DB 241 dvlcpggrylkagcaelpncidldgfcacagfelgkrgscvsgsgoptlgng 300
 OY 301 VPTRRPATATSPVQRTWPRIWDEKLGCTPLVPEODNSVTSIPEIPIKWSQSTNTSLQW 360
 DB 301 vptrrpatatspvqrtwprirwdeklgctplvpodnsvtsipeipikwsqstntslq 360
 OY 361 SIQASKATITPSSGVISKFNSTSSATPOADSSAVFIVSTRVAVLILMTVIGL 420
 DB 361 siqaeskatitpssgviskfnstssatpoadssavfivstrvavllilmtvlg 420
 OY 421 VKICHPSSOPRKESMPGLESDPEPALGSSSAHCTNNGVGVGCDLDRRAEGALL 480
 DB 421 VKICHPSSOPRKESMPGLESDPEPALGSSSAHCTNNGVGVGCDLDRRAEGALL 480
 OY 481 AESPLGSSDA 490
 DB 481 aesplgssda 490
 RESULT 7
 AAB41847
 ID AAB41847 standard; Protein: 331 AA.
 XX
 AC AAB41847;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORF1611 polypeptide sequence SEQ ID NO:3222.
 XX
 KW Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
 KW vulnary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antihemetic; antihypertoid;
 KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

| | | | | | |
|----|--|---------|---------------------|--------|--------------------|
| | Query Match | 60.48; | Score 296; | DB 21; | Length 331; |
| | Best Local Similarity: | 100.0%; | Pred. No. 1,4e-284; | | |
| | Matches 296; Conservative | 0; | Mismatches | 0; | Indels 10; Gaps 0; |
| OY | 147 QATGVPEAPAKMKRMCHLRANGTYLCKYOFEVYLCPAPRGAASNLSTVPAPOLSAALDES | 206 | | | |
| Dd | | | | | |
| | 33 .qatgvpeapawkcmrthlrangylykgyfevylcpaprpgaaasnlstvpqrqlnsaalids | 92 | | | |
| OY | 207 PPTGEVSVALRGGLPITSVCIADELGAWKDLSDGVLCPCPGRYLRAGKCAELPNCLIDL | 266 | | | |
| Dd | | | | | |
| | 93 pptgevsalrcgglplstvciaadelgaagtwklsdgvlpcpcgrylragkaaelpnclidl | 152 | | | |
| OY | 267 GGRCACATGFELGKDGRSCVSSGGCOTTLGGTGVPPRRPATTSVDPQRTPIIVNDEK | 326 | | | |
| Dd | | | | | |
| | 153 ggacacaagfelgkdgrscvssggcgqpfllgltvpprrpatatsvdpqrtwpiivndek | 212 | | | |

| Accession | Gene | Protein | Seq ID | Seq ID NO. |
|-----------|---|---------|--------|------------|
| QY 327 | GFEPVLPEDNNTSTSPETPRNGSGSTMTSLQMSLQAEKATTPSGSVTSKFNSTSS | 386 | | |
| Db 213 | Igepvlpednntstspetprngsgstmtslqmslqaeaktltpsgsvtskfnstss | 386 | | |
| QY 387 | ATPOADSSAAVFIFVSTAVVVLVTMTVYLGKLCFHSPSSOPPKESMGPRG | 442 | | |
| Db 273 | atpoadssaaavfifvstavvvlvtmtvygkclcfhspssoppkessmgppg | 328 | | |
| RESULT 8 | | | | |
| AA91410 | | | | |
| ID | AA91410 standard. Protein. 344 AA. | | | |
| AC | | | | |
| XX | AA91410; | | | |
| DT | | | | |
| DE | 29-JUN-2000 (first entry) | | | |
| XX | | | | |
| DE | Human secreted protein sequence encoded by gene 4 SEQ ID NO:131. | | | |
| XX | | | | |
| KW | Human: secreted protein; diagnosis; neuroprotective; nootropic; | | | |
| KW | neuroleptic; anti-magic; cerebroprotective; immunomodulatory; | | | |
| KW | anti-microbial; cardiac; cytosolic; anti-inflammatory; haemostatic; | | | |
| KW | anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy; | | | |
| KW | neural; reproductive; immune disorder; immunodeficiency; infection; | | | |
| KW | aneurysm; demyelinating disease; autoimmunity; cancer; inflammation; | | | |
| KW | aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease; | | | |
| KW | Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis | | | |
| KW | ischaemia; mania; dementia; obsessive compulsive disorder; | | | |
| KW | viral prophylaxis; developmental disorder; sexually-linked disorder; | | | |
| KW | cardiovascular disorder; food additive; preservative; chromosome 14. | | | |
| XX | | | | |
| OS | Homo sapiens. | | | |
| XX | | | | |
| XX | W0200011014-1. | | | |
| XX | | | | |
| PD | 02-MAR-2000. | | | |
| XX | | | | |
| XX | 24-AUG-1999; 99MO-US19330. | | | |
| XX | | | | |
| XX | 25-AUG-1998; 98US-0097917. | | | |
| XX | | | | |
| PR | 31-AUG-1998; 98US-0098634. | | | |
| XX | | | | |
| PA | (HUMA-) HUMAN GENOME SCI INC. | | | |
| PI | | | | |
| PI | Moore PA, Ruben SM, Olsef HS, Shi Y, Rosen CA, Florence KA; | | | |
| PI | Soppet DR, Laflaur DW, Endress GA, Edner R, Komatsoulis G; | | | |
| PI | Duan RD; | | | |
| XX | | | | |
| DR | WPI: 2000-224656/19. | | | |
| XX | | | | |
| PT | Novel secreted proteins and corresponding DNA molecules that can be | | | |
| PT | used to prevent, treat and diagnose disease in humans, for example, | | | |
| PT | Alzheimer's, cancer, and immune disorders - | | | |
| XX | | | | |
| PS | Disclosure: Page 391-392; 416pp. English. | | | |
| XX | | | | |
| CC | The polynucleotide sequences given in AAA26281 to AAA26336 encode the | | | |
| CC | human secreted proteins given in AA91346 to AA91449. The human secreted | | | |
| CC | proteins can have activities based on the tissues and cells they are | | | |
| CC | expressed in. Examples of the activities are: neuroprotective; nootropic; | | | |
| CC | neuroleptic; anti-magic; cerebroprotective; immunomodulatory; | | | |
| CC | anti-microbial; cardiac; cytosolic; anti-inflammatory; haemostatic; | | | |
| CC | anticonvulsant; and vasotropic. The polynucleotides and proteins may be | | | |
| CC | used to prevent, treat or ameliorate a medical condition, e.g. by protein | | | |
| CC | or gene therapy. Conditions treatable by the proteins of the invention | | | |
| CC | include neural, reproductive, or immune disorders, especially | | | |
| CC | immunodeficiency, infection, lymphomas, demyelinating diseases, | | | |
| CC | auto-immunities, cancer, general microbial infection, inflammation, | | | |
| CC | aneurysms and haemorrhages. Specific examples include: Alzheimer's | | | |
| CC | disease; Parkinson's; Huntington's; Tourette syndrome; multiple | | | |
| CC | sclerosis; meningitis; ischaemia; prostate cancer; mania; dementia; | | | |
| CC | obsessive compulsive disorder and viral prophylaxis. | | | |

CC and proteins can also be used in the detection of disorders associated
 CC with the function of the protein, for example, the detection of
 CC developmental disorders, sexually-linked disorders, or disorders of the
 CC cardiovascular system. They may also be used as food additives or
 CC 'preservatives'. AAA26272 to AAA26280 and AAY91345 are sequences used in
 CC the exemplification of the present invention.

XX Sequence 344 AA;

Query Match 55.9%; Score 274; DB 21; Length 344;
 Best Local Similarity 100.0%; Pred. No. 9.4e-263;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 YLCKYQFVLCPPAPGASNLSTYRAPQLHSAALDFSPGTEVSALRGQLPISVTCTA 228
 Db 23 YLCKYQFVLCPPAPGASNLSTYRAPQLHSAALDFSPGTEVSALRGQLPISVTCTA 82
 OY 229 DEIGARMKLSGDVLCPCPGRLRAGKCAELPNCUDLGGFACGATGFEKDGRCVY 288
 Db 83 deigerwklisgdlvclpcpgrylrqgkcaelpnclddlygfecacatgfejqkdgrrscvt 142
 OY 289 SGEQPTLGTGCVPRRRPATSPVORTWPIRVDKLGEPPLVEQDNSTVSIPEIR 348
 Db 143 sgegptlgtgcvprrrpatatpvtwpiirvdeklgetplvpeqdnstvtspelpir 202
 OY 349 WGSOSTMSTLQMSLQESKATIPSGSVISKFNSTSSATPOAPDSSSAVVEIFVSTAV 408
 Db 203 wsgostmstlqmslqeskatitpsgsviskfnstssatpafssavvelfvstlav 262
 OY 409 VLVITMTVLGIVKICFHESSPSOPRKESMGP 442
 Db 263 vlvitmtvlglvkvicfhespsoprkessmgypp 296

RESULT 9

AAB80228
 ID AAB80228 standard; Protein: 428 AA.

XX AAB80228;

DT 24-APR-2001 (first entry)

XX Human PRO269 protein.

XX Human; PRO: dermatological; antipsoriatic; cytostatic; antinflammatory;

KW antiParkinsonian nootropic; neuroprotective; vulnerary; cardiac;

KW antiangiogenic; vasotropic; antihistaminic; antirheumatic; cancer;

KW antiarthritic; antiinfectility; antidiabetic; antiviral; diabetes;

KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;

XX schaeemia; inflammation.

XX Homo sapiens.

XX WO200104311-A1.

XX 18-JAN-2001.

XX 22-FEB-2000; 2000WO-US04414.

XX 07-JUL-1999; 99US-0143048.

XX 26-JUL-1999; 99US-0145638.

XX 28-JUL-1999; 99US-0146222.

XX 08-SEP-1999; 99WO-US20594.

XX 13-SEP-1999; 99WO-US20944.

XX 15-SEP-1999; 99WO-US21050.

XX 15-SEP-1999; 99WO-US21547.

XX 05-OCT-1999; 99WO-US23089.

XX 29-NOV-1999; 99WO-US28214.

XX 30-NOV-1999; 99WO-US28313.

XX 16-DEC-1999; 99WO-US30095.

XX 20-DEC-1999; 99WO-US30911.

XX 20-DEC-1999; 99WO-US30999.

PR 05-JAN-2000; 99WO-US00219.

XX (GETH) GENENTECH INC.

PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;

PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;

PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kijavlin IJ;

PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;

PI Williams PM, Wood WI;

PI WPI: 2001-081051/09.

DR N-PSDB: AAF72289.

PT Sixty one nucleic acids encoding PRO polypeptides which are useful in

PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung

PT squamous cell carcinoma) and neurodegenerative diseases (e.g.

PT Alzheimer's disease) -

XX Claim 1: Fig 36; 393pp; English.

XX The present sequence is one of sixty one novel secreted and

XX transmembrane PRO polypeptides. The PRO polypeptides are

XX useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung

XX squamous cell carcinoma), gastrointestinal disorders (e.g. lung

XX enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,

XX Parkinson's disease), wound repair, cardiovascular disorders (e.g.

XX endometrial bleeding angiogenesis, ischaemias such as coronary

XX ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,

XX rheumatoid arthritis, multiple sclerosis), infertility, AIDS and

XX diabetes and retinal disorders such as retinitis pigmentosa.

XX The PRO nucleic acids have applications in molecular biology, including

XX use as hybridization probes, and in chromosome and gene mapping.

XX Sequence 428 AA;

Query Match 49.4%; Score 242; DB 22; Length 428;

Best Local Similarity 100.0%; Pred. No. 5.8e-231;
 Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 RYLKAGKCAELPNCUDLGGFACGATGFEKDGRCVYSGEGPTLGTGVPTRRPPA 308

Db 187 rylrkgkcaelpnclddlygfecacatgfejqkdgrrscvtsggptlgtgvprrppa 246

OY 309 TATSPVORTWPIRVDKLGEPPLVEQDNSTVSIPEIRPMGOSTMSTLQMSLQESKA 368

Db 247 tatsvpqrtwpiirvdeklgetplvpeqdnstvtspelpirwsgtmslqmslqeska 306

OY 369 TTTPSGSVYSKFNSTSSATPOAPDSSSAVVEIFVSTAVVLVITMTVLGIVKICFHE 428

Db 307 tltpsgsvyskfnstssatpafssavvelfvstlavvilitmtvlglvkvicfhes 366

OY 429 PSSOPRKESMGPGLSPDEPAALGSSAHCTNNCKVGDCLDRBAGALAESPLSS 488

Db 367 pssoprkessmgypplesdepaaligssahcnnckvkgcdclrdraegaalaaesplss 426

OY 489 DA 490

Db 427 da 428

RESULT 10

AAY95031
 ID AAY95031 standard; Protein: 175 AA.

XX AAY95031;

DT 19-JUN-2000 (first entry)

XX Human clone vpl5.1 insertional variant ORF, SEQ ID NO:131.

XX Human; secreted protein; cancer; tumour; cardiovascular disorder;

KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;

and proteins can also be used in the detection of disorders associated with

and proteins can also be used in the detection of disorders associated

RESULT 11
AAV91349 ID AAV91349 standard; Protein, 189 AA.
AAV91349; ID
29-JUN-2000 (first entry)
Human secreted protein sequence encoded by gene 4 SEQ ID NO:70.
Human: secreted protein; diagnosis; neuroprotective; nootropic;
neuroleptic; antimanic; cerebroprotective; immunomodulatory;
anti-microbial; cardiant; cyostatic; antiinflammatory; haemostatic;
anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
neural; reproductive; immune disorder; immunodeficiency; infection;
lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
Huntington's disease; Tourette syndrome; multiple sclerosis; meningo-
ischemia; mania; dementia; obsessive compulsive disorder;
viral prophylaxis; developmental disorder; sexually-linked disorder;
cardiovascular disorder; food additive; preservative; chromosome 14.
Homo sapiens.
WO200011014-A1.
02-MAR-2000.
24-AUG-1999; 99WO-US19330.
25-AUG-1998; 98US-0097917.
31-AUG-1998; 98US-0098634.
(HOMA-) HUMAN GENOME SCI INC.
Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
Soppet DI, Laflleur DW, Endress GA, Ebner R, Komatsu G;
Duan RD;
WPI: 2000-224656/19.
N-PDB: AAA26284.
Novel secreted proteins and corresponding DNA molecules that can be
used to prevent, treat or ameliorate a disease in humans, for example,
Alzheimer's, cancer, and immune disorders -
Claim 11; Page 357; 416pp; English.
The polynucleotide sequences given in AAA26281 to AAA26336 encode the
human secreted proteins given in AAV91346 to AAV91449. The human secreted
proteins can have activities based on the tissues and cells they are
expressed in. Examples of the activities are: neuroprotective; nootropic;
neuroleptic; antimanic; cerebroprotective; immunomodulatory;
anti-microbial; cardiant; cyostatic; antiinflammatory; nootropic;
anticonvulsant; and vasotropic. The polynucleotides and proteins may be
used to prevent, treat or ameliorate a medical condition, e.g. by protein
or gene therapy. Conditions treatable by the proteins of the invention
include neural, reproductive, or immune disorders, especially
immunodeficiency, infection, lymphomas, demyelinating diseases,
auto-immunities, cancer, general microbial infection, inflammation,
aneurysms and haemorrhages. Specific examples include: Alzheimer's
disease; Parkinson's; Huntington's; Tourette syndrome; multiple
sclerosis; meningitis; ischemia; prostate cancer; mania; dementia;
obsessive compulsive disorder and viral prophylaxis. The polynucleotides
and proteins can also be used in the detection of disorders associated

CC with the function of the protein, for example, the detection of
 CC developmental disorders, sexually-linked disorders, or disorders of the
 CC cardiovascular system. They may also be used as food additives or
 CC preservatives. AAA26272 to AAA26280 and AAY91345 are sequences used in
 CC the exemplification of the present invention.

XX Sequence 189 AA:

Query Match 17.1%; Score 84; DB 21; Length 189;
 Best Local Similarity 100.0%; Pred. No. 11e-74;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPAFLCLLWQALMPGCGEHPADRAGCSAGCAYSLHATMKROAEACILRGCA 60

DB 1 mrpafalcllwqalmpgpggghpdragscagacyslhatmkrgaaeacilrgga 60

OY 61 LSTVAGAEIRVALLLRAGPG 84

DB 61 lstrvagaelrvalllrpgpg 84

RESULT 12

AA91409

ID AAY91409 standard; Protein: 123 AA.

AA91409;

29-JUN-2000 (first entry)

Human secreted protein sequence encoded by gene 4 SEQ ID NO:130.
 Human; secreted protein; diagnosis; neuroprotective; nootropic;
 neuroleptic; antimanic; cerebroprotective; immunomodulatory;
 anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
 anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
 neural; reproductive; immune disorder; immunodeficiency; infection;
 lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
 aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
 Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
 ischaemia; mania; dementia; obsessive compulsive disorder;
 viral prophylaxis; developmental disorder; sexually-linked disorder;
 cardiovascular disorder; food additive; preservative; chromosome 14.

XX Homo sapiens.

OS WO200011014-A1.

XX 02-MAR-2000.

XX 24-AUG-1999; 99WO-US19330.

XX 25-AUG-1998; 98US-0097917.

XX 31-AUG-1998; 98US-0098634.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
 PI Soppet DR, Lafleur DW, Endress GA, Edner R, Komatsoulis G;
 PI Duan RD;

XX WPI: 2000-224656/19.

XX Novel secreted proteins and corresponding DNA molecules that can be
 PT used to prevent, treat and diagnose disease in humans, for example,
 PT Alzheimer's, cancer, and immune disorders -

XX Disclosure: Page 391; 416pp; English.

XX The polynucleotide sequences given in AAA26281 to AAA26336 encode the
 CC human secreted proteins given in AAY91346 to AAY91449. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: neuroprotective; nootropic;

CC neuroleptic; antimanic; cerebroprotective; immunomodulatory;
 CC anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
 CC anticonvulsant; and vasotropic. The polynucleotides and proteins may be
 CC used to prevent, treat or ameliorate a medical condition, e.g. by protein
 CC or gene therapy. Conditions treatable by the proteins of the invention
 CC include neural, reproductive, or immune disorders, especially
 CC immunodeficiency, infection, lymphomas, demyelinating diseases,
 CC auto-immunities, cancer, general microbial infection, inflammation,
 CC aneurysms and haemorrhages. Specific examples include: Alzheimer's
 CC disease; Parkinson's; Huntington's; Tourette syndrome; multiple
 CC sclerosis; meningitis; ischaemia; prostate cancer; mania; dementia;
 CC obsessive compulsive disorder and viral prophylaxis. The polynucleotides
 CC and proteins can also be used in the detection of disorders associated
 CC with the function of the protein, for example, the detection of
 CC developmental disorders, sexually-linked disorders, or disorders of the
 CC cardiovascular system. They may also be used as food additives or
 CC preservatives. AAA26272 to AAA26280 and AAY91345 are sequences used in
 CC the exemplification of the present invention.

XX Sequence 123 AA:

Query Match 12.9%; Score 63; DB 21; Length 123;
 Best Local Similarity 100.0%; Pred. No. 4.7e-54;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 EHPTADRAGCSAGCAYSLHATMKROAEACILRGALSTVAGAEIRVALLLRAGP 81

DB 1 ehptadrgcsagacyslhatmkrgaaeacilrggalstvrageelrvalllrpg 60

OY 82 GPG 84

DB 61 gpg 63

RESULT 13

ABG08184

ID ABG08184 standard; Protein: 186 AA.

ABG08184;

13-FEB-2002 (first entry)

Novel human diagnostic protein #8175.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HISE-) HYSEQ INC.

XX Dimaac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX N-PSDB: AAS72371.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID NO 38543; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG030377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

SO Sequence 186 AA;
 Query Match 1.88; Score 9; DB 22; Length 186;
 Best Local Similarity 100.08; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 78 RAGPQGGC 86;
 Db 139 ragpqp99 147

RESULT 14

AAU54703 standard; Protein: 79 AA.

AC AAU54703;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #15599.

KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 dermatological; osteopathic; neuroprotectant.

KW Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

PA Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A.

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71

DR N-PSDB: AAS5956.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris

PS Example 1; SEQ ID No 15898; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

SO Sequence 79 AA;

Query Match 1.68; Score 8; DB 22; Length 79;
 Best Local Similarity 100.08; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 ALERRRRSH 102
 Db 20 alerrrrsh 27

RESULT 15

AAU43842 standard; Protein: 123 AA.

AC AAU43842;

XX 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #4738.

KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 dermatological; osteopathic; neuroprotectant.

KW Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

PA Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A.

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71

DR N-PSDB: AAS59521.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:37:56 ; Search time 30.68 Seconds

(without alignments)
2762.957 Million cell updates/sec

Title: US-09-902-713b-96

Perfect score: 490
Sequence: 1 MRPALCLLMQALMPGPG.....LRDRAGALLAESPLGSSDA 490

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Wo size: 0
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

SPTREMBL.19:**
1: sp_archaea:**
2: sp_bacteria:**
3: sp_fungi:**
4: sp_human:**
5: sp_invertebrate:**
6: sp_mammal:**
7: sp_mhc:**
8: sp_organelle:**
9: sp_phage:**
10: sp_plant:**
11: sp_rodent:**
12: sp_virus:**
13: sp_vertebrate:**
14: sp_unclassified:**
15: sp_virus:**
16: sp_bacteriopl:**
17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 158 | 32.2 | 212 | 4 Q9P096 | Q9P096 homo sapien |
| 2 | 24 | 4.9 | 459 | 11 Q9DC55 | Q9DC55 mus musculu |
| 3 | 24 | 4.9 | 459 | 11 Q9D624 | Q9D624 mus musculu |
| 4 | 24 | 4.9 | 459 | 11 Q9CXA8 | Q9CXA8 mus musculu |
| 5 | 8 | 1.6 | 187 | 2 Q9E0U8 | Q9E0U8 salmonella |
| 6 | 8 | 1.6 | 222 | 2 Q9ADC6 | Q9ADC6 streptomyc |
| 7 | 8 | 1.6 | 371 | 5 Q9VZS1 | Q9VZS1 petromyzo |
| 8 | 8 | 1.6 | 438 | 13 Q9OZM7 | Q9OZM7 dirosophila |
| 9 | 8 | 1.6 | 464 | 11 Q9JLF3 | Q9JLF3 petromyzo |
| 10 | 8 | 1.6 | 491 | 4 Q9NS54 | Q9NS54 mus musculu |
| 11 | 8 | 1.6 | 498 | 2 Q9JAS3 | Q9JAS3 homo sapien |
| 12 | 8 | 1.6 | 498 | 2 Q9JAS3 | Q9JAS3 pseudomonas |
| 13 | 8 | 1.6 | 730 | 12 Q9QZU3 | Q9QZU3 mus musculu |
| 14 | 8 | 1.6 | 730 | 12 Q9QZU3 | Q9QZU3 bovine herp |
| 15 | 8 | 1.6 | 1343 | 12 Q9G635 | Q9G635 bovine herp |
| 16 | 8 | 1.6 | 1816 | 16 Q9KBB4 | Q9KBB4 bacillus ha |

| | | | | | |
|----|---|-----|-----|-----------|---------------------|
| 17 | 7 | 1.4 | 72 | 12 Q69143 | Q69143 human herpe |
| 18 | 7 | 1.4 | 76 | 2 Q9AJX7 | Q9AJX7 streptomyc |
| 19 | 7 | 1.4 | 80 | 2 Q9EMN2 | Q9EMN2 streptomyc |
| 20 | 7 | 1.4 | 88 | 17 Q970G5 | Q970G5 sulfolobus |
| 21 | 7 | 1.4 | 92 | 2 Q9X7P3 | Q9X7P3 streptomyc |
| 22 | 7 | 1.4 | 97 | 10 Q49401 | Q49401 arabidopsis |
| 23 | 7 | 1.4 | 102 | 15 Q91931 | Q91931 human immun |
| 24 | 7 | 1.4 | 102 | 15 Q9YV65 | Q9YV65 human immun |
| 25 | 7 | 1.4 | 103 | 15 Q91930 | Q91930 human immun |
| 26 | 7 | 1.4 | 109 | 2 Q9ZNB8 | Q9ZNB8 nitrosomona |
| 27 | 7 | 1.4 | 115 | 6 Q9Z921 | Q9Z921 sus scrofa |
| 28 | 7 | 1.4 | 127 | 4 Q9B282 | Q9B282 homo sapien |
| 29 | 7 | 1.4 | 130 | 2 Q9ZB13 | Q9ZB13 streptomyc |
| 30 | 7 | 1.4 | 131 | 4 Q9H5Y0 | Q9H5Y0 homo sapien |
| 31 | 7 | 1.4 | 136 | 2 Q9X8S4 | Q9X8S4 streptomyc |
| 32 | 7 | 1.4 | 139 | 2 Q9XAP5 | Q9XAP5 streptomyc |
| 33 | 7 | 1.4 | 150 | 4 Q9S923 | Q9S923 homo sapien |
| 34 | 7 | 1.4 | 151 | 2 Q9K5J4 | Q9K5J4 buchnera ap |
| 35 | 7 | 1.4 | 151 | 2 Q9K5J3 | Q9K5J3 buchnera ap |
| 36 | 7 | 1.4 | 151 | 2 Q9K5J2 | Q9K5J2 buchnera ap |
| 37 | 7 | 1.4 | 151 | 2 Q9K5J1 | Q9K5J1 buchnera ap |
| 38 | 7 | 1.4 | 151 | 2 Q9K5J0 | Q9K5J0 buchnera ap |
| 39 | 7 | 1.4 | 151 | 2 Q9K5I9 | Q9K5I9 buchnera ap |
| 40 | 7 | 1.4 | 151 | 2 Q9K5I8 | Q9K5I8 buchnera ap |
| 41 | 7 | 1.4 | 151 | 2 Q9K5I7 | Q9K5I7 buchnera ap |
| 42 | 7 | 1.4 | 151 | 2 Q9K5I6 | Q9K5I6 buchnera ap |
| 43 | 7 | 1.4 | 151 | 2 Q9K5I5 | Q9K5I5 buchnera ap |
| 44 | 7 | 1.4 | 151 | 2 Q9K5I4 | Q9K5I4 buchnera ap |
| 45 | 7 | 1.4 | 151 | 2 Q9K5I3 | Q9K5I3 buchnera ap |
| 46 | 7 | 1.4 | 151 | 2 Q9K5I2 | Q9K5I2 buchnera ap |
| 47 | 7 | 1.4 | 151 | 2 Q9K5I1 | Q9K5I1 buchnera ap |
| 48 | 7 | 1.4 | 151 | 2 Q9K5C8 | Q9K5C8 secondary e |
| 49 | 7 | 1.4 | 151 | 2 Q9K5C7 | Q9K5C7 secondary e |
| 50 | 7 | 1.4 | 151 | 2 Q9K5C6 | Q9K5C6 secondary e |
| 51 | 7 | 1.4 | 151 | 2 Q93V14 | Q93V14 buchnera ap |
| 52 | 7 | 1.4 | 151 | 11 Q91VW0 | Q91VW0 mus musculu |
| 53 | 7 | 1.4 | 152 | 12 Q9DUD0 | Q9DUD0 tt virus. o |
| 54 | 7 | 1.4 | 155 | 4 Q96G14 | Q96G14 homo sapien |
| 55 | 7 | 1.4 | 155 | 16 Q920X5 | Q920X5 rhizobium m |
| 56 | 7 | 1.4 | 161 | 5 Q9GPG6 | Q9GPG6 trypanosoma |
| 57 | 7 | 1.4 | 161 | 5 Q9U0U0 | Q9U0U0 leishmania |
| 58 | 7 | 1.4 | 166 | 2 Q87504 | Q87504 escherichia |
| 59 | 7 | 1.4 | 172 | 14 Q991W0 | Q991W0 uncultured |
| 60 | 7 | 1.4 | 173 | 2 P72384 | P72384 streptomyc |
| 61 | 7 | 1.4 | 180 | 15 Q91872 | Q91872 human immun |
| 62 | 7 | 1.4 | 190 | 16 Q9RMB3 | Q9RMB3 deinoococcus |
| 63 | 7 | 1.4 | 191 | 16 Q9A6B2 | Q9A6B2 caulobacter |
| 64 | 7 | 1.4 | 191 | 17 Q9YF50 | Q9YF50 aeropyrum p |
| 65 | 7 | 1.4 | 193 | 4 Q92605 | Q92605 homo sapien |
| 66 | 7 | 1.4 | 195 | 11 Q60539 | Q60539 mesocricetu |
| 67 | 7 | 1.4 | 198 | 4 Q12870 | Q12870 homo sapien |
| 68 | 7 | 1.4 | 199 | 4 Q9N001 | Q9N001 homo sapien |
| 69 | 7 | 1.4 | 203 | 4 Q13222 | Q13222 homo sapien |
| 70 | 7 | 1.4 | 207 | 10 Q9S903 | Q9S903 arabidopsis |
| 71 | 7 | 1.4 | 212 | 17 Q9Y9C9 | Q9Y9C9 aeropyrum p |
| 72 | 7 | 1.4 | 213 | 16 Q9RY03 | Q9RY03 deinoococcus |
| 73 | 7 | 1.4 | 214 | 11 Q9D0P3 | Q9D0P3 mus musculu |
| 74 | 7 | 1.4 | 216 | 5 P90877 | P90877 caenorhabdi |
| 75 | 7 | 1.4 | 217 | 11 Q9C8S4 | Q9C8S4 mus musculu |
| 76 | 7 | 1.4 | 218 | 16 Q985B5 | Q985B5 rhizobium l |
| 77 | 7 | 1.4 | 223 | 16 Q9RUD8 | Q9RUD8 deinoococcus |
| 78 | 7 | 1.4 | 228 | 11 Q70473 | Q70473 cricetus |
| 79 | 7 | 1.4 | 230 | 10 Q94J39 | Q94J39 oryza sativ |
| 80 | 7 | 1.4 | 237 | 2 Q9RGE7 | Q9RGE7 clostridium |
| 81 | 7 | 1.4 | 242 | 2 Q938N5 | Q938N5 streptococc |
| 82 | 7 | 1.4 | 243 | 16 Q9RX66 | Q9RX66 deinoococcus |
| 83 | 7 | 1.4 | 246 | 2 Q9K431 | Q9K431 streptomyc |
| 84 | 7 | 1.4 | 249 | 16 Q9AAW0 | Q9AAW0 caulobacter |
| 85 | 7 | 1.4 | 252 | 11 Q9CT53 | Q9CT53 mus musculu |
| 86 | 7 | 1.4 | 252 | 11 Q9C0I5 | Q9C0I5 mus musculu |
| 87 | 7 | 1.4 | 253 | 16 Q9RYD6 | Q9RYD6 deinoococcus |
| 88 | 7 | 1.4 | 254 | 5 Q27230 | Q27230 taenia ovis |
| 89 | 7 | 1.4 | | | |

DE HSPC285 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_taxonomy9606;
RN (1)
RN SEQUENCE FROM N.A.
RP
RP TISSUE=BLOOD;
RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.O., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT "Human partial CDS from cd34+ stem cells.";
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF161403; AAF28963.1; -
FT NONIER
SQ
SEQUENCE 212 AA; 22078 MW; 12E1194E7A8F75D1 CRC64;

| Query Match | 32.2% | Score 158 | DB 4 | Length 212 |
|---|---|---|------|------------|
| Best Local Similarity 100.0% | | | | |
| Matches 158, Conservative 0, Mismatches 0, Indels 0, Gaps 0 | | | | |
| QY | 285 | SCVYSGGCGPTLGGTGVPRRPATATSPVORTPIRVDEKLEETPLVPEONSVTSIP | 344 | |
| Db | 52 | SCVYSGGCGPTLGGTGVPRRPATATSPVORTPIRVDEKLEETPLVPEONSVTSIP | 111 | |
| QY | 345 | EIPFMGSGQSMSTLQMSLQAEKATITTSYGVSISKFNSTSSATPAQFDSSAVFIFFS | 404 | |
| Db | 112 | EIPFMGSGQSMSTLQMSLQAEKATITTSYGVSISKFNSTSSATPAQFDSSAVFIFFS | 171 | |
| QY | 405 | TAVVLYVILITMTVLGLVKLCFHEPSSQPRRESMGPPG | 442 | |
| Db | 172 | TAVVLYVILITMTVLGLVKLCFHEPSSQPRRESMGPPG | 209 | |
| RESULT | 2 | | | |
| ID | Q9DC55 | PRELIMINARY; | PRT; | 459 AA. |
| AC | Q9DC55: | | | |
| DT | 01-JUN-2001 (TREMBLrel, 17, Created) | | | |
| DT | 01-JUN-2001 (TREMBLrel, 17, Last sequence update) | | | |
| DT | 01-DEC-2001 (TREMBLrel, 19, Last annotation update) | | | |
| GN | 1200003C23RIK PROTEIN. | | | |
| DE | 1200003C23RIK | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; | | | |
| OX | NCBI_TaxID=10090; | | | |
| RP | 111 | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=C57BL/6J; TISSUE=LUNG; | | | |
| RC | MEDLINE=21085660; PubMed=11217851. | | | |
| RA | Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., | | | |
| RA | Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., | | | |
| RA | Aizawa K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamana K. I., | | | |
| RA | Saito T., Okazaki Y., Gojohori T., Bono H., Kusakawa T., Saito R., | | | |
| RA | Kodota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., | | | |
| RA | Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H., | | | |
| RA | Knehl P., Lewis S., Matsuo Y., Nikardo I., Pesole G., Quackenbush J., | | | |
| RA | Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., | | | |
| RA | Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G., | | | |
| RA | Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., | | | |
| RA | Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., | | | |
| RA | Gastlisch S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H., | | | |
| RA | Lyons P., Marchionni L., Mashima J., Mazzarelli J., Membrants P., | | | |
| RA | Nordone P., Ring B., Schoenbach C., Seya T., Shibata Y., Storch K.F., | | | |
| RA | Sasaki H., Sato K., Schoenbach C., Rodriguez I., Sakamoto N., | | | |
| RA | Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L., | | | |
| RA | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S., | | | |
| RA | Hayashizaki Y. | | | |
| RT | Functional annotation of a full-length mouse cDNA collection. " | | | |
| RT | Nature 409:685-690(2001). | | | |
| DR | EMBL: AK004557; BAB23370.1; - | | | |
| DR | HSP; P00743; IICF. | | | |

DR MGD:1914114; 1200003C23R1K.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_c.
 DR SMART: SM00034; CLECF; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 DR SEQUENCE 459 AA; 49053 MW; 6CB25D048A456A7B CRC64;

Query Match 4.9%; Score 24; DB 11; Length 459;
 Best Local Similarity 100.0%; Pred. No. 3.4e-15;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GYLCKYQFVLCPPAPRGASNLIS 191
 DB 168 GYLCKYQFVLCPPAPRGASNLIS 191

AC 09D624 PRELIMINARY; PRT; 459 AA.
 AC 09D624;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE 1200003C23R1K PROTEIN.
 GN 1200003C23R1K
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Arakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirimi L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilmink L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK014681; BAB29502.1; -
 DR HSSP: P00743; 1CCF.
 DR MGD: MGI:1914114; 1200003C23R1K.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_c.
 DR SMART: SM00034; CLECF; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 DR SEQUENCE 459 AA; 49065 MW; A257B38DA598EC8A CRC64;

Query Match 4.9%; Score 24; DB 11; Length 459;

Best Local Similarity 100.0%; Pred. No. 3.4e-15;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GYLCKYQFVLCPPAPRGASNLIS 191
 DB 168 GYLCKYQFVLCPPAPRGASNLIS 191

AC 09CXAB PRELIMINARY; PRT; 459 AA.
 AC 09CXAB;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE 1200003C23R1K PROTEIN.
 GN 1200003C23R1K
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC LUNG;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Arakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirimi L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilmink L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK014682; BAB31209.1; -
 DR HSSP: P00743; 1CCF.
 DR MGD: MGI:1914114; 1200003C23R1K.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001304; lectin_c.
 DR SMART: SM00034; CLECF; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 DR SEQUENCE 459 AA; 49150 MW; A649730F15132DCD CRC64;

Query Match 4.9%; Score 24; DB 11; Length 459;
 Best Local Similarity 100.0%; Pred. No. 3.4e-15;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GYLCKYQFVLCPPAPRGASNLIS 191
 DB 168 GYLCKYQFVLCPPAPRGASNLIS 191

AC 09EUB8 PRELIMINARY; PRT; 187 AA.
 AC 09EUB8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)

```

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
GN ORE27 (CONJUGATIVE TRANSFER: REGULATION)
OS Salmomella enterica subsp. enterica serovar Choleraesuis, and
OS Salmomella typhimurium LT2
OC Plasmid 50k virulence plasmid, and plasmid pSLT
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmomella
OX NCBI_TaxID=119912, 99287;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium LT2; STRAIN=SGSC1412; PLASMID=PSLT;
RC PLASMID=50k VIRULENCE PLASMID;
RX MEDLINE=21153631; PubMed=11254626;
RA Handed T., Okada N., Nakazawa N., Kawakami T., Danbara H.,
RT "Complete DNA Sequence and Comparative Analysis of the 50-kilobase
RT virulence plasmid of Salmomella enterica Serovar Choleraesuis."
P Infect. Immun. 69:2612-2620(2001).
[2]
SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium LT2; STRAIN=SGSC1412; PLASMID=PSLT;
RA WashU:
RT "The Salmomella typhimurium Genome Sequencing Project."
RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium LT2; STRAIN=SGSC1412; PLASMID=PSLT;
RA McClelland M., Sanderson K., Spieth J., Clifton S., Latreille P.,
RA Courtney L., All J., Dane M., Du F., Hou S., Layman D., Leonard S.,
RA Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E.,
RA Sun H., Florea L., Miller W., Porwollik S., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB040415; BAB20534.1;
DR HSSP: P29367; IDVO.
KW Plasmid.
SQ SEQUENCE 187 AA; 21133 MW; D6E2D389CAEE9EB8 CRC64;

Query Match
Best Local Similarity: 1.6%; Score 8; DB 2; Length 187;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 AELRAVLA 75
      |||||
      177 AELRAVLA 184

```

```

RESULT 6
ID 09ADC6; PRELIMINARY; PRT; 222 AA.
AC 09ADC6;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE SP2A2 PROTEIN
GN SP2A2
OS Streptomyces coelicolor.
OC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.,
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinsahl H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL589707; CAC33905.1;
SQ SEQUENCE 222 AA; 23251 MW; B74F3719B4BF310A CRC64;

Query Match
Best Local Similarity: 1.6%; Score 8; DB 2; Length 222;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 444 ESDEPPAA 451
      |||||
      141 ESDEPPAA 148

RESULT 7
ID 090Y00; PRELIMINARY; PRT; 284 AA.
AC 090Y00;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE STEROID RECEPTOR (FRAGMENT)
OS Petromyzon marinus (Sea Lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RA Escriva H., Manzoni L., Youson J., Laudet V.;
RT "Analysis of cyclostone genes reveals waves of gene duplication during
RT early vertebrate evolution."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF16879; AAL26247.1;
KW Receptor.
FT NON_TER 1 1
FT NON_TER 284 284
SQ SEQUENCE 284 AA; 32284 MW; 6E6DC2PC07F27C27 CRC64;

Query Match
Best Local Similarity: 1.6%; Score 8; DB 13; Length 284;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 TTSSAPQ 390
      |||||
      63 TTSSAPQ 70

RESULT 8
ID 09VZS1; PRELIMINARY; PRT; 371 AA.
AC 09VZS1;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE CG17746 PROTEIN
GN CG17746
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:37:01 ; Search time 16.1 Seconds

(without alignments)
743.388 Million cell updates/sec

Title: US-09-902-713B-96

Perfect score: 490

Sequence: 1 MRPAFLCLIMQALWPGG.....LRDRAGALLAESPLGSSDA 490

Scoring table:
OLIGO
Gapop 60.0, Gapext 60.0
231628 seqs, 24425594 residues

Search: 231628 seqs, 24425594 residues

Wop size: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database: Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 8 | 1.6 | 234 | 1 | US-08-299-567-5 |
| 2 | 8 | 1.6 | 238 | 1 | US-08-240-124-2 |
| 3 | 8 | 1.6 | 238 | 1 | US-08-453-943-2 |
| 4 | 8 | 1.6 | 238 | 2 | US-09-057-121-2 |
| 5 | 8 | 1.6 | 238 | 4 | US-09-358-734-2 |
| 6 | 7 | 1.4 | 330 | 2 | US-08-712-948-1 |
| 7 | 7 | 1.4 | 333 | 2 | US-08-712-948-1 |
| 8 | 7 | 1.4 | 834 | 4 | US-09-143-571-29 |
| 9 | 7 | 1.4 | 1239 | 4 | US-08-937-931-2 |
| 10 | 7 | 1.4 | 1239 | 4 | US-09-285-502-2 |
| 11 | 7 | 1.4 | 1239 | 4 | US-09-709-126-2 |
| 12 | 7 | 1.4 | 2089 | 4 | US-08-418-893D-23 |
| 13 | 7 | 1.4 | 2089 | 1 | US-08-418-893D-20 |
| 14 | 6 | 1.2 | 14 | 1 | US-08-445-745-12 |
| 15 | 6 | 1.2 | 14 | 1 | US-08-447-010-21 |
| 16 | 6 | 1.2 | 15 | 1 | US-07-859-291C-37 |
| 17 | 6 | 1.2 | 15 | 2 | US-08-124-981A-30 |
| 18 | 6 | 1.2 | 18 | 3 | US-09-037-190-26 |
| 19 | 6 | 1.2 | 18 | 3 | US-09-037-192-26 |
| 20 | 6 | 1.2 | 18 | 4 | US-09-037-143-26 |
| 21 | 6 | 1.2 | 18 | 4 | US-09-049-691-26 |
| 22 | 6 | 1.2 | 18 | 4 | US-08-260-174-26 |
| 23 | 6 | 1.2 | 18 | 4 | US-09-338-128A-26 |
| 24 | 6 | 1.2 | 20 | 2 | US-09-232-346-26 |
| 25 | 6 | 1.2 | 20 | 2 | US-08-124-981A-26 |
| 26 | 6 | 1.2 | 20 | 2 | US-09-010-928B-7 |
| 27 | 6 | 1.2 | 23 | 3 | US-08-855-531D-38 |

| | | | | | | |
|-----|---|-----|-----|---|--------------------|--------------------|
| 28 | 6 | 1.2 | 23 | 4 | US-08-855-526B-38 | Sequence 38, Appl |
| 29 | 6 | 1.2 | 24 | 4 | US-09-439-897-62 | Sequence 62, Appl |
| 30 | 6 | 1.2 | 25 | 2 | US-08-902-516-4 | Sequence 4, Appl |
| 31 | 6 | 1.2 | 29 | 3 | US-08-488-551B-636 | Sequence 636, App |
| 32 | 6 | 1.2 | 29 | 6 | 5221620-7 | Patent No. 5221620 |
| 33 | 6 | 1.2 | 31 | 4 | US-09-071-224-36 | Sequence 36, Appl |
| 34 | 6 | 1.2 | 34 | 3 | US-08-388-353-636 | Sequence 636, App |
| 35 | 6 | 1.2 | 36 | 1 | US-08-428-248-4 | Sequence 4, Appl |
| 36 | 6 | 1.2 | 39 | 4 | US-09-220-528-49 | Sequence 49, Appl |
| 37 | 6 | 1.2 | 44 | 3 | US-09-053-197A-66 | Sequence 66, Appl |
| 38 | 6 | 1.2 | 44 | 4 | US-09-085-761A-71 | Sequence 71, Appl |
| 39 | 6 | 1.2 | 55 | 1 | US-08-442-063A-60 | Sequence 60, Appl |
| 40 | 6 | 1.2 | 61 | 1 | US-08-358-160-147 | Sequence 147, App |
| 41 | 6 | 1.2 | 62 | 4 | US-09-402-668-12 | Sequence 12, App |
| 42 | 6 | 1.2 | 73 | 4 | US-09-564-805-213 | Sequence 213, App |
| 43 | 6 | 1.2 | 78 | 6 | 5514590-7 | Patent No. 5514590 |
| 44 | 6 | 1.2 | 91 | 2 | US-08-341-843B-31 | Sequence 31, Appl |
| 45 | 6 | 1.2 | 91 | 2 | US-08-427-497E-36 | Sequence 36, Appl |
| 46 | 6 | 1.2 | 93 | 2 | US-08-327-362-1 | Sequence 1, Appl |
| 47 | 6 | 1.2 | 93 | 4 | US-09-158-565-1 | Sequence 1, Appl |
| 48 | 6 | 1.2 | 105 | 3 | US-08-946-329A-103 | Sequence 103, App |
| 49 | 6 | 1.2 | 111 | 4 | US-09-220-528-53 | Sequence 53, Appl |
| 50 | 6 | 1.2 | 114 | 4 | US-09-230-637-28 | Sequence 28, Appl |
| 51 | 6 | 1.2 | 115 | 2 | US-07-903-029-4 | Sequence 4, Appl |
| 52 | 6 | 1.2 | 115 | 2 | US-07-903-029-5 | Sequence 5, Appl |
| 53 | 6 | 1.2 | 116 | 1 | US-07-634-278-57 | Sequence 57, Appl |
| 54 | 6 | 1.2 | 116 | 1 | US-07-634-278-73 | Sequence 73, Appl |
| 55 | 6 | 1.2 | 116 | 1 | US-08-477-728-57 | Sequence 57, Appl |
| 56 | 6 | 1.2 | 116 | 1 | US-08-477-728-73 | Sequence 73, Appl |
| 57 | 6 | 1.2 | 116 | 1 | US-08-474-040-57 | Sequence 57, Appl |
| 58 | 6 | 1.2 | 116 | 1 | US-08-474-040-73 | Sequence 73, Appl |
| 59 | 6 | 1.2 | 116 | 1 | US-08-487-200-57 | Sequence 57, Appl |
| 60 | 6 | 1.2 | 116 | 1 | US-08-487-200-73 | Sequence 73, Appl |
| 61 | 6 | 1.2 | 116 | 4 | US-08-484-537-57 | Sequence 57, Appl |
| 62 | 6 | 1.2 | 119 | 4 | US-08-484-537-73 | Sequence 73, Appl |
| 63 | 6 | 1.2 | 119 | 1 | US-08-442-542-45 | Sequence 45, Appl |
| 64 | 6 | 1.2 | 119 | 1 | US-08-285-936-2 | Sequence 2, Appl |
| 65 | 6 | 1.2 | 119 | 1 | US-08-285-936-6 | Sequence 6, Appl |
| 66 | 6 | 1.2 | 119 | 1 | US-08-285-936-8 | Sequence 8, Appl |
| 67 | 6 | 1.2 | 119 | 1 | US-08-285-936-10 | Sequence 10, Appl |
| 68 | 6 | 1.2 | 119 | 1 | US-08-285-936-12 | Sequence 12, Appl |
| 69 | 6 | 1.2 | 119 | 1 | US-08-469-486-49 | Sequence 49, Appl |
| 70 | 6 | 1.2 | 119 | 1 | US-08-487-860-2 | Sequence 2, Appl |
| 71 | 6 | 1.2 | 119 | 1 | US-08-487-860-6 | Sequence 6, Appl |
| 72 | 6 | 1.2 | 119 | 1 | US-08-487-860-8 | Sequence 8, Appl |
| 73 | 6 | 1.2 | 119 | 1 | US-08-487-860-10 | Sequence 10, Appl |
| 74 | 6 | 1.2 | 119 | 1 | US-08-487-860-12 | Sequence 12, Appl |
| 75 | 6 | 1.2 | 119 | 1 | US-08-487-860-57 | Sequence 57, Appl |
| 76 | 6 | 1.2 | 119 | 2 | US-08-469-658-49 | Sequence 49, Appl |
| 77 | 6 | 1.2 | 119 | 3 | US-08-765-469-45 | Sequence 45, Appl |
| 78 | 6 | 1.2 | 120 | 1 | US-08-446-919A-6 | Sequence 6, Appl |
| 79 | 6 | 1.2 | 123 | 3 | US-08-938-548B-10 | Sequence 10, Appl |
| 80 | 6 | 1.2 | 123 | 3 | US-08-939-093A-10 | Sequence 10, Appl |
| 81 | 6 | 1.2 | 129 | 2 | US-08-530-165-5 | Sequence 5, Appl |
| 82 | 6 | 1.2 | 130 | 3 | US-08-938-548B-6 | Sequence 6, Appl |
| 83 | 6 | 1.2 | 130 | 4 | US-08-939-093A-6 | Sequence 6, Appl |
| 84 | 6 | 1.2 | 131 | 1 | US-07-779-706A-2 | Sequence 2, Appl |
| 85 | 6 | 1.2 | 131 | 1 | US-08-104-503-2 | Sequence 2, Appl |
| 86 | 6 | 1.2 | 137 | 2 | US-08-621-751A-4 | Sequence 4, Appl |
| 87 | 6 | 1.2 | 137 | 2 | US-08-621-751A-8 | Sequence 8, Appl |
| 88 | 6 | 1.2 | 141 | 2 | US-08-765-783A-29 | Sequence 29, Appl |
| 89 | 6 | 1.2 | 141 | 2 | US-08-921-100-29 | Sequence 29, Appl |
| 90 | 6 | 1.2 | 141 | 3 | US-08-880-142-29 | Sequence 29, Appl |
| 91 | 6 | 1.2 | 141 | 3 | US-08-902-201-29 | Sequence 29, Appl |
| 92 | 6 | 1.2 | 141 | 4 | US-09-416-557-29 | Sequence 29, Appl |
| 93 | 6 | 1.2 | 146 | 1 | US-08-054-480-2 | Sequence 2, Appl |
| 94 | 6 | 1.2 | 164 | 3 | US-09-009-913-11 | Sequence 11, Appl |
| 95 | 6 | 1.2 | 170 | 2 | US-08-477-451-35 | Sequence 35, Appl |
| 96 | 6 | 1.2 | 171 | 3 | US-09-011-735-2 | Sequence 2, Appl |
| 97 | 6 | 1.2 | 174 | 4 | US-08-858-207A-377 | Sequence 377, App |
| 98 | 6 | 1.2 | 183 | 1 | US-08-447-010-2 | Sequence 2, Appl |
| 99 | 6 | 1.2 | 183 | 1 | US-08-447-010-4 | Sequence 4, Appl |
| 100 | 6 | 1.2 | 188 | 5 | PCT-US94-14277-5 | Sequence 5, Appl |

Sequence 15, Appl
Sequence 15, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 9, Appl
Sequence 20, Appl
Sequence 20, Appl
Patent No. 5,187,262
Sequence 5, Appl
Sequence 7, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 16, Appl
Sequence 29, Appl
Sequence 16, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 12, Appl
Sequence 31, Appl
Sequence 19, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 7, Appl
Sequence 305, Appl
Sequence 7, Appl
Sequence 198, Appl
Sequence 2, Appl
Sequence 26, Appl
Sequence 1, Appl
Sequence 16, Appl
Sequence 8, Appl
Sequence 62, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 6, Appl
Sequence 8, Appl
Sequence 9, Appl

| | | | | |
|-----------------------|---------|----------------|-------|---------------|
| Query Match | 1.68; | Score 8; | DB 1; | Length 234; |
| Best Local Similarity | 100.08; | Pred. No. 8.5; | | |
| Matches | 8; | Conservative | 0; | Mismatches 0. |

RESULT 2
US-08-240-124-2
Sequence 2, Application: US/08240124
Patent No. 5516658

GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,124
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SESEE, KATHYNN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:36:21 ; Search time 20.98 Seconds

(without alignments)
2244.222 Million cell updates/sec

Title: US-09-902-713b-96

Perfect score: 490

Sequence: 1 MRPAFALCLIMQALMPGGG.....LRDRAEGALLAESPLGSSDA 490

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database: PIR_71:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 8 | 1.6 | 238 | 2 I38849 | LEER-3 - human |
| 2 | 8 | 1.6 | 299 | 2 JC4374 | steroid uptake prot |
| 3 | 8 | 1.6 | 491 | 2 T46915 | hypothetical prote |
| 4 | 8 | 1.6 | 1041 | 2 S55862 | probable membrane |
| 5 | 8 | 1.6 | 1816 | 2 F83901 | hypothetical prote |
| 6 | 8 | 1.4 | 55 | 2 AH2074 | hypothetical prote |
| 7 | 7 | 1.4 | 76 | 2 AG3389 | hypothetical prote |
| 8 | 7 | 1.4 | 92 | 2 T35262 | hypothetical prote |
| 9 | 7 | 1.4 | 97 | 2 G85212 | hypothetical prote |
| 10 | 7 | 1.4 | 113 | 2 T05023 | hypothetical prote |
| 11 | 7 | 1.4 | 130 | 2 T35978 | hypothetical prote |
| 12 | 7 | 1.4 | 133 | 2 S58123 | chlorodoxin (clone |
| 13 | 7 | 1.4 | 134 | 2 T36547 | gasirin-releasing |
| 14 | 7 | 1.4 | 136 | 2 T36547 | hypothetical prote |
| 15 | 7 | 1.4 | 139 | 2 T34633 | hypothetical prote |
| 16 | 7 | 1.4 | 151 | 2 S37647 | H+-transporting AT |
| 17 | 7 | 1.4 | 190 | 2 H75478 | conserved hypothet |
| 18 | 7 | 1.4 | 191 | 2 F72731 | hypothetical prote |
| 19 | 7 | 1.4 | 191 | 2 E87519 | Thyl/PlpI family P |
| 20 | 7 | 1.4 | 198 | 2 A57717 | transcription fact |
| 21 | 7 | 1.4 | 207 | 2 E86429 | hypothetical prote |
| 22 | 7 | 1.4 | 212 | 2 C72464 | hypothetical prote |
| 23 | 7 | 1.4 | 213 | 2 F75585 | urease accessory p |
| 24 | 7 | 1.4 | 213 | 2 B84333 | isopenenyl pyroph |
| 25 | 7 | 1.4 | 216 | 2 T22314 | hypothetical prote |
| 26 | 7 | 1.4 | 223 | 2 G75394 | conserved hypothet |
| 27 | 7 | 1.4 | 238 | 2 S03906 | 45k antigen - shee |
| 28 | 7 | 1.4 | 243 | 2 E75517 | hypothetical prote |
| 29 | 7 | 1.4 | 249 | 2 B87309 | hypothetical prote |

| | | | | | |
|-----|---|-----|-----|----------|--------------------|
| 30 | 7 | 1.4 | 253 | 2 E75570 | gldb protein - Del |
| 31 | 7 | 1.4 | 259 | 2 T49596 | related to H+-tran |
| 32 | 7 | 1.4 | 264 | 2 E83284 | probable enoyl-CoA |
| 33 | 7 | 1.4 | 268 | 2 A72552 | hypothetical prote |
| 34 | 7 | 1.4 | 268 | 2 A12212 | hypothetical prote |
| 35 | 7 | 1.4 | 269 | 2 T36001 | probable integral |
| 36 | 7 | 1.4 | 270 | 2 S65739 | basigin precursor |
| 37 | 7 | 1.4 | 278 | 2 C86421 | FIN18.17 protein - |
| 38 | 7 | 1.4 | 290 | 2 G71304 | probable translati |
| 39 | 7 | 1.4 | 290 | 2 A69779 | conserved hypothet |
| 40 | 7 | 1.4 | 293 | 2 T09758 | myb-related protei |
| 41 | 7 | 1.4 | 302 | 2 H82638 | hypothetical prote |
| 42 | 7 | 1.4 | 309 | 2 E64224 | yabc protein homol |
| 43 | 7 | 1.4 | 312 | 2 T43853 | conserved hypothet |
| 44 | 7 | 1.4 | 312 | 2 AH2389 | hypothetical prote |
| 45 | 7 | 1.4 | 313 | 2 D86829 | ribose ABC transpo |
| 46 | 7 | 1.4 | 313 | 2 C69231 | conserved hypothet |
| 47 | 7 | 1.4 | 318 | 1 S30914 | 2-desacetyl-2-hydr |
| 48 | 7 | 1.4 | 318 | 1 T50752 | alcohol dehydrogen |
| 49 | 7 | 1.4 | 325 | 1 JN0629 | alcohol dehydrogen |
| 50 | 7 | 1.4 | 325 | 2 A33851 | transcription fact |
| 51 | 7 | 1.4 | 328 | 2 S35336 | homeotic protein H |
| 52 | 7 | 1.4 | 330 | 2 A40855 | conserved hypothet |
| 53 | 7 | 1.4 | 339 | 2 F83026 | udp-glucose 4-epim |
| 54 | 7 | 1.4 | 344 | 2 G71858 | type III secretion |
| 55 | 7 | 1.4 | 345 | 2 D91201 | glutamate decarbox |
| 56 | 7 | 1.4 | 345 | 2 H86047 | hypothetical prote |
| 57 | 7 | 1.4 | 355 | 2 D84192 | hypothetical prote |
| 58 | 7 | 1.4 | 360 | 2 D65045 | hypothetical prote |
| 59 | 7 | 1.4 | 360 | 2 H91068 | hypothetical prote |
| 60 | 7 | 1.4 | 360 | 2 B85813 | hypothetical prote |
| 61 | 7 | 1.4 | 376 | 2 S69529 | thiamin biosynthes |
| 62 | 7 | 1.4 | 377 | 2 AC0933 | hypothetical prote |
| 63 | 7 | 1.4 | 383 | 2 AD0504 | probable acyl-CoA |
| 64 | 7 | 1.4 | 401 | 2 G83197 | probable Mn transp |
| 65 | 7 | 1.4 | 411 | 2 B97309 | hypothetical prote |
| 66 | 7 | 1.4 | 412 | 2 AB2471 | glycine hydroxymet |
| 67 | 7 | 1.4 | 418 | 2 H81383 | hypothetical prote |
| 68 | 7 | 1.4 | 414 | 2 T29470 | hypothetical prote |
| 69 | 7 | 1.4 | 423 | 2 H83603 | protein Y48A6B.10 |
| 70 | 7 | 1.4 | 433 | 2 C88594 | transcription fact |
| 71 | 7 | 1.4 | 439 | 2 I57561 | hypothetical prote |
| 72 | 7 | 1.4 | 446 | 2 T26988 | competence factor |
| 73 | 7 | 1.4 | 449 | 2 G95004 | serine hydroxymet |
| 74 | 7 | 1.4 | 451 | 2 D69356 | flavonol 3-O-gluc |
| 75 | 7 | 1.4 | 455 | 1 XUBHFG | ATP synthase beta |
| 76 | 7 | 1.4 | 458 | 2 C82952 | H+-transporting tw |
| 77 | 7 | 1.4 | 460 | 1 PMECB | H+-transporting AT |
| 78 | 7 | 1.4 | 460 | 2 B91213 | ATP synthase beta |
| 79 | 7 | 1.4 | 460 | 2 C86059 | H+-transporting AT |
| 80 | 7 | 1.4 | 460 | 2 AF0954 | ATP synthase beta |
| 81 | 7 | 1.4 | 460 | 2 AE0500 | H+-transporting AT |
| 82 | 7 | 1.4 | 461 | 2 G86594 | lipamide dehydrog |
| 83 | 7 | 1.4 | 461 | 2 C72031 | 2-oxo acid dehydro |
| 84 | 7 | 1.4 | 461 | 2 C64796 | dec protein |
| 85 | 7 | 1.4 | 461 | 2 D90711 | c4-dicarboxylate a |
| 86 | 7 | 1.4 | 461 | 2 H85561 | transport of dicar |
| 87 | 7 | 1.4 | 461 | 2 AF0579 | H+-transporting AT |
| 88 | 7 | 1.4 | 465 | 2 H84930 | H+-transporting AT |
| 89 | 7 | 1.4 | 465 | 2 C81970 | ATP synthase F1, b |
| 90 | 7 | 1.4 | 465 | 2 G81024 | ATP synthase, beta |
| 91 | 7 | 1.4 | 466 | 2 G82715 | H+-transporting AT |
| 92 | 7 | 1.4 | 467 | 2 S06082 | ATP synthase F1, b |
| 93 | 7 | 1.4 | 467 | 2 F82036 | hypothetical prote |
| 94 | 7 | 1.4 | 467 | 2 T34874 | H+-transporting AT |
| 95 | 7 | 1.4 | 468 | 2 D64071 | hypothetical prote |
| 96 | 7 | 1.4 | 468 | 2 T22598 | H+-transporting AT |
| 97 | 7 | 1.4 | 477 | 2 S23257 | alpha-amylase (EC |
| 98 | 7 | 1.4 | 482 | 1 S40887 | RVS167 protein - y |
| 99 | 7 | 1.4 | 488 | 2 C75488 | argininosuccinate |
| 100 | 7 | 1.4 | 493 | 2 T37606 | probable succinate |
| 101 | 7 | 1.4 | 502 | 2 S36494 | E2 protein - human |
| 102 | 7 | 1.4 | 506 | 2 F83547 | sodium/proline sym |

gldb protein - Del
related to H+-tran
probable enoyl-CoA
hypothetical prote
hypothetical prote
probable integral
basigin precursor
FIN18.17 protein -
probable translati
conserved hypothet
myb-related protei
hypothetical prote
yabc protein homol
conserved hypothet
hypothetical prote
ribose ABC transpo
conserved hypothet
2-desacetyl-2-hydr
alcohol dehydrogen
alcohol dehydrogen
transcription fact
homeotic protein H
conserved hypothet
udp-glucose 4-epim
type III secretion
glutamate decarbox
hypothetical prote
hypothetical prote
hypothetical prote
thiamin biosynthes
hypothetical prote
probable acyl-CoA
probable Mn transp
hypothetical prote
glycine hydroxymet
hypothetical prote
protein Y48A6B.10
transcription fact
hypothetical prote
competence factor
serine hydroxymet
flavonol 3-O-gluc
ATP synthase beta
H+-transporting tw
H+-transporting AT
ATP synthase beta
H+-transporting AT
ATP synthase beta
lipamide dehydrog
2-oxo acid dehydro
dec protein
c4-dicarboxylate a
transport of dicar
H+-transporting AT
H+-transporting AT
ATP synthase F1, b
ATP synthase, beta
H+-transporting AT
ATP synthase F1, b
hypothetical prote
H+-transporting AT
hypothetical prote
alpha-amylase (EC
RVS167 protein - y
argininosuccinate
probable succinate
E2 protein - human
sodium/proline sym

103 7 1.4 507 2 G01614 zinc finger protei
104 7 1.4 509 2 S76731 hypothetical prote
105 7 1.4 510 1 W2BE47 protein-serine/thr
106 7 1.4 540 2 G50068 choline transpor
107 7 1.4 549 2 T23526 hypothetical prote
108 7 1.4 561 2 C68837 nodulin-like prote
109 7 1.4 563 2 T36704 probable dehydrog
110 7 1.4 567 2 A86262 hypothetical prote
111 7 1.4 579 2 T24880 hypothetical prote
112 7 1.4 584 2 T29469 beta-fructofuran
113 7 1.4 592 2 S5661 hypothetical prote
114 7 1.4 605 2 AH1150 ABC transporter, A
115 7 1.4 605 2 A11509 ABC transporter, A
116 7 1.4 608 2 T35755 probable branched
117 7 1.4 611 2 D70928 threonyl-tRNA synt
118 7 1.4 640 2 F83303 hypothetical prote
119 7 1.4 647 2 S67651 hypothetical prote
120 7 1.4 661 2 I52603 MSP1 protein - mou
121 7 1.4 674 2 D72329 hypothetical prote
122 7 1.4 678 2 B70913 probable penicilli
123 7 1.4 690 2 A83441 potassium transpor
124 7 1.4 702 1 J00401 triphosphan synth
125 7 1.4 724 2 D48941 ATP-dependent tran
126 7 1.4 738 2 D70680 hypothetical prote
127 7 1.4 749 2 I37356 epithelial microtu
128 7 1.4 797 2 G83570 probable two-compo
129 7 1.4 826 2 A60385 REDI protein - yea
130 7 1.4 827 1 C0BYD1 trac protein homol
131 7 1.4 856 2 T31234 collagen alpha 2 c
132 7 1.4 991 2 A35763 hypothetical prote
133 7 1.4 1051 2 F96831 probable transcrip
134 7 1.4 1076 2 D71476 probable disintegr
135 7 1.4 1239 2 T13809 hypothetical prote
136 7 1.4 1312 2 T32738 DNA-directed RNA p
137 7 1.4 1487 2 T02850 hypothetical prote
138 7 1.4 1530 1 S01393 hypothetical prote
139 7 1.4 1642 2 T19130 F1003.10 protein -
140 7 1.4 1744 2 A48757 acetyl-CoA carboxy
141 7 1.4 2089 1 F86161 gellan lyase (EC 4
142 7 1.4 2475 1 W2BE56 367k tegument prot
143 7 1.4 3421 1 LDL-receptor-relat
144 7 1.4 4151 2 T13734 LDL-receptor-relat
145 7 1.4 4753 2 T14593 syringomycin synth
146 7 1.4 9376 2 S40139 T-cell receptor J-
147 7 1.4 24 2 T04371 thumectin-like pro
148 6 1.2 26 2 T35208 hypothetical prote
149 6 1.2 45 2
150 6 1.2

ALIGNMENTS

RESULT 1
138849
LEKR-3 - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
C:Accession: I38849
R:Kozlowsky, C.J.; Maraskovsky, E.; McGrew, J.T.; Vandenbos, T.; Teepe, M.; Lyman, S.D.;
Oncogene 10, 299-306, 1995
A:Title: Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs encod
A:Reference number: I38849; MUID:95140419
A:Accession: I38849
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-238 <RES>
C:Genetics:
A:Cross-references: EMBL:U14187, NID:9642832; PID:AA050078.1; PID:9642833
A:Gene: GDB:EP163
A:Cross-references: GDB:438336; OMIM:601381
A:Map position: 1q21-1q22
A:Family: axon guidance signal protein

Query Match 1.6%; Score 8; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 AGPGGGG 86
DB 75 AGPGGGG 82

RESULT 2

JC4374
sterol uptake protein 1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G1828; protein YGL162w; SUT1 protein
C:Species: Saccharomyces cerevisiae
C:Date: 10-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 15-Sep-2000
C:Accession: JC4374; S60424; S64178
R:Bourot, S.; Karst, F.
Gene 165, 97-102, 1995
A:Title: Isolation and characterization of the Saccharomyces cerevisiae SUT1 gene inv
A:Reference number: JC4374; MUID:96084960
A:Accession: JC4374
A:Molecule type: DNA
A:Residues: 1-299 <BOU>
A:Cross-references: EMBL:X77766; NID:g1183994; PID:CAAS4806.1; PID:g1183995
R:James, C.M.; Indge, K.J.; Oliver, S.G.
Yeast 11, 1413-1415, 1995
A:Title: DNA sequence analysis of a 35 kb segment from Saccharomyces cerevisiae chrom
A:Reference number: S60424; MUID:96158061
A:Accession: S60424
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-299 <JAM>
A:Cross-references: EMBL:Z48618
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995
R:James, C.M.; Indge, K.J.; Oliver, S.G.
Submitted to the Protein Sequence Database, May 1996
A:Reference number: S64165
A:Accession: S64178
A:Molecule type: DNA
A:Residues: 1-299 <JAM>
A:Cross-references: EMBL:Z72684; NID:g1322757; PID:e243599; PID:g1322758; MIPS:YGL162
A:Experimental source: strain S288C
C:Comment: This protein is a member of the hypoxic gene family, and is rich in serine
C:Genetics:
A:Gene: SGP:SUT1
A:Cross-references: SGP:S0003130; MIPS:YGL162w
A:Map position: 7L
C:Superfamily: Saccharomyces cerevisiae hypothetical protein YPR009w
F:133-151/Region: serine/threonine-rich

Query Match 1.6%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 381 NSTSSAT 388
DB 138 NSTSSAT 145

RESULT 3
T46915
hypothetical protein DKFZP762A227.1 - human
C:Species: Homo sapiens (man)
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: T46915
R:Ottewill, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
Submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24136
A:Accession: T46915

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:39:11 ; Search time 13.5 Seconds
(without alignments)
1405.376 Million cell updates/sec

Title: US-09-902-713b-96

Perfect score: 490

Sequence: 1 MKPAPALCLIMQALMPGPGC.....LRDRAEGALLAESPLGSSDA 490

Scoring table: OLIGO

Gapop 60.0 ; Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database: SwissProt_40:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 8 | 1.6 | 238 | 1 | ERR3_HUMAN |
| 2 | 8 | 1.6 | 299 | 1 | SOT1_YEAST |
| 3 | 8 | 1.6 | 1041 | 1 | EGT2_YEAST |
| 4 | 7 | 1.4 | 133 | 1 | THH2_ARATH |
| 5 | 7 | 1.4 | 134 | 1 | GRE_SHEEP |
| 6 | 7 | 1.4 | 155 | 1 | MOAE_RHIME |
| 7 | 7 | 1.4 | 192 | 1 | Y4GC_RHISN |
| 8 | 7 | 1.4 | 213 | 1 | IDT_HALNT |
| 9 | 7 | 1.4 | 240 | 1 | 45KD_TAEQV |
| 10 | 7 | 1.4 | 270 | 1 | BASI_RABIT |
| 11 | 7 | 1.4 | 290 | 1 | EFPS_TREPA |
| 12 | 7 | 1.4 | 309 | 1 | Y222_MYCGE |
| 13 | 7 | 1.4 | 318 | 1 | BCHC_RHOSH |
| 14 | 7 | 1.4 | 324 | 1 | ALDX_HUMAN |
| 15 | 7 | 1.4 | 324 | 1 | ALDX_RAT |
| 16 | 7 | 1.4 | 325 | 1 | YCAT_ECOLI |
| 17 | 7 | 1.4 | 328 | 1 | CEBB_CHICK |
| 18 | 7 | 1.4 | 330 | 1 | THH1_HUMAN |
| 19 | 7 | 1.4 | 376 | 1 | HTH1_SALTY |
| 20 | 7 | 1.4 | 376 | 1 | YOZ3_BHPI |
| 21 | 7 | 1.4 | 414 | 1 | GLYA_CAME |
| 22 | 7 | 1.4 | 416 | 1 | PROA_LEPIN |
| 23 | 7 | 1.4 | 431 | 1 | SMA6_CHICK |
| 24 | 7 | 1.4 | 449 | 1 | COMB_STRPN |
| 25 | 7 | 1.4 | 451 | 1 | GLYA_ARCFU |
| 26 | 7 | 1.4 | 455 | 1 | UFQG_HORVU |
| 27 | 7 | 1.4 | 457 | 1 | ATPB_HAEIN |
| 28 | 7 | 1.4 | 457 | 1 | ATPB_PASWU |
| 29 | 7 | 1.4 | 459 | 1 | ATPB_ECOLI |
| 30 | 7 | 1.4 | 461 | 1 | DCUC_ECOLI |
| 31 | 7 | 1.4 | 461 | 1 | DLDH_CHLPP |
| 32 | 7 | 1.4 | 463 | 1 | ATPB_BURCE |
| 33 | 7 | 1.4 | 463 | 1 | ATPB_BURCE |

| | | | | | |
|-----|---|-----|------|---|------------|
| 34 | 7 | 1.4 | 465 | 1 | ATPB_BUCAL |
| 35 | 7 | 1.4 | 466 | 1 | ATPB_BUCAP |
| 36 | 7 | 1.4 | 466 | 1 | ATPB_VIBAL |
| 37 | 7 | 1.4 | 466 | 1 | ATPB_VIBCH |
| 38 | 7 | 1.4 | 468 | 1 | ATPB_THREE |
| 39 | 7 | 1.4 | 482 | 1 | R167_YEAST |
| 40 | 7 | 1.4 | 507 | 1 | VE2_HPV25 |
| 41 | 7 | 1.4 | 507 | 1 | MKR3_HUMAN |
| 42 | 7 | 1.4 | 510 | 1 | KR2_VZVD |
| 43 | 7 | 1.4 | 592 | 1 | INV2_DAUCA |
| 44 | 7 | 1.4 | 611 | 1 | VT13_MYCTU |
| 45 | 7 | 1.4 | 625 | 1 | DUS8_HUMAN |
| 46 | 7 | 1.4 | 625 | 1 | AMT_ALTRA |
| 47 | 7 | 1.4 | 669 | 1 | NAD6_MYCTU |
| 48 | 7 | 1.4 | 679 | 1 | FZD8_MOUSE |
| 49 | 7 | 1.4 | 685 | 1 | ATRB_PSEAE |
| 50 | 7 | 1.4 | 690 | 1 | TRP_COPCI |
| 51 | 7 | 1.4 | 702 | 1 | TRP_COPCI |
| 52 | 7 | 1.4 | 724 | 1 | PEPD_PEDOC |
| 53 | 7 | 1.4 | 827 | 1 | ADOB_MOUSE |
| 54 | 7 | 1.4 | 827 | 1 | CADG_HUMAN |
| 55 | 7 | 1.4 | 829 | 1 | MSH5_HUMAN |
| 56 | 7 | 1.4 | 834 | 1 | Y419_HUMAN |
| 57 | 7 | 1.4 | 991 | 1 | TLR9_HUMAN |
| 58 | 7 | 1.4 | 1032 | 1 | TLR9_MOUSE |
| 59 | 7 | 1.4 | 1083 | 1 | T2D3_HUMAN |
| 60 | 7 | 1.4 | 1250 | 1 | TP3A_DROME |
| 61 | 7 | 1.4 | 1330 | 1 | PURA_HUMAN |
| 62 | 7 | 1.4 | 1530 | 1 | RPCI_TRYBB |
| 63 | 7 | 1.4 | 1544 | 1 | TUSP_HUMAN |
| 64 | 7 | 1.4 | 1547 | 1 | TUSP_MOUSE |
| 65 | 7 | 1.4 | 3421 | 1 | TRGU_HSVB |
| 66 | 7 | 1.4 | 4753 | 1 | LRP_CABEL |
| 67 | 7 | 1.2 | 14 | 1 | TAT_HV123 |
| 68 | 6 | 1.2 | 29 | 1 | LHA2_ECTHA |
| 69 | 6 | 1.2 | 47 | 1 | YK02_CAEEL |
| 70 | 6 | 1.2 | 50 | 1 | PSBU_SYNGU |
| 71 | 6 | 1.2 | 65 | 1 | SECE_THEMA |
| 72 | 6 | 1.2 | 71 | 1 | HPIS_PARSP |
| 73 | 6 | 1.2 | 73 | 1 | BB13_SCHCO |
| 74 | 6 | 1.2 | 74 | 1 | Y4HR_RHISN |
| 75 | 6 | 1.2 | 76 | 1 | NEUG_BOVIN |
| 76 | 6 | 1.2 | 78 | 1 | NEUG_CAPHI |
| 77 | 6 | 1.2 | 78 | 1 | NEUG_HUMAN |
| 78 | 6 | 1.2 | 78 | 1 | NEUG_RAT |
| 79 | 6 | 1.2 | 83 | 1 | RL27_THDMA |
| 80 | 6 | 1.2 | 86 | 1 | ISK4_MOUSE |
| 81 | 6 | 1.2 | 86 | 1 | TAT_HV126 |
| 82 | 6 | 1.2 | 86 | 1 | Y97_ADE07 |
| 83 | 6 | 1.2 | 92 | 1 | Y36_ARCFU |
| 84 | 6 | 1.2 | 92 | 1 | SRG1_MOUSE |
| 85 | 6 | 1.2 | 98 | 1 | SRG1_MOUSE |
| 86 | 6 | 1.2 | 98 | 1 | SRG1_MOUSE |
| 87 | 6 | 1.2 | 98 | 1 | SRG1_MOUSE |
| 88 | 6 | 1.2 | 98 | 1 | SRG1_MOUSE |
| 89 | 6 | 1.2 | 101 | 1 | TAT_HV126 |
| 90 | 6 | 1.2 | 101 | 1 | TAT_HV126 |
| 91 | 6 | 1.2 | 103 | 1 | EMF1_CHICK |
| 92 | 6 | 1.2 | 103 | 1 | Y688_HAEIN |
| 93 | 6 | 1.2 | 104 | 1 | BCHA_ERESP |
| 94 | 6 | 1.2 | 107 | 1 | SSIS_STRPR |
| 95 | 6 | 1.2 | 110 | 1 | VLI_FVUL |
| 96 | 6 | 1.2 | 110 | 1 | YH07_YEAST |
| 97 | 6 | 1.2 | 112 | 1 | DAD1_DROME |
| 98 | 6 | 1.2 | 113 | 1 | FRT2_HUMAN |
| 99 | 6 | 1.2 | 115 | 1 | GUAN_HUMAN |
| 100 | 6 | 1.2 | 115 | 1 | HV44_MOUSE |
| 101 | 6 | 1.2 | 116 | 1 | ND3M_ONCKE |
| 102 | 6 | 1.2 | 116 | 1 | ND3M_ONCKE |
| 103 | 6 | 1.2 | 116 | 1 | ND3M_ONCKE |
| 104 | 6 | 1.2 | 116 | 1 | ND3M_ONCKE |
| 105 | 6 | 1.2 | 116 | 1 | ND3M_ONCKE |
| 106 | 6 | 1.2 | 116 | 1 | ND3M_ONCKE |

| | |
|--------|--------------|
| P57124 | buchnera ap |
| Q07232 | buchnera ap |
| P12986 | vibrio algi |
| Q9knh5 | vibrio chol |
| P41168 | thiobacilli |
| P39743 | saccharomyc |
| P36787 | human papil |
| Q13064 | homo sapien |
| P09286 | varicella-z |
| Q39692 | daucus caro |
| Q10830 | mycobacteri |
| Q13202 | homo sapien |
| P29957 | alteromonas |
| P71911 | mycobacteri |
| Q61081 | mus musculu |
| P57688 | pseudomonas |
| P16578 | coprinus ci |
| P36497 | pediococcus |
| Q05910 | mus musculu |
| P14291 | saccharomyc |
| Q75309 | homo sapien |
| Q43196 | homo sapien |
| Q43303 | homo sapien |
| Q9nr96 | homo sapien |
| Q9eqn3 | mus musculu |
| Q00268 | homo sapien |
| Q9ng98 | drosophila |
| O15067 | homo sapien |
| P08968 | trypanosoma |
| Q9nr14 | homo sapien |
| Q9nj15 | mus musculu |
| P28955 | equine herp |
| Q04833 | caenorhabdi |
| P12511 | human immun |
| P12510 | human immun |
| P80101 | ecoli |
| P34297 | caenorhabdi |
| P56152 | synechococc |
| P35874 | thermotoga |
| P00264 | paracoccus |
| P55482 | rhizobium s |
| P35722 | bos taurus |
| P54877 | capra hircu |
| Q92686 | homo sapien |
| Q04940 | rattus norv |
| Q9x197 | thermotoga |
| Q35679 | mus musculu |
| P12506 | human immun |
| P04509 | human immun |
| P05668 | human adeno |
| Q28637 | archaeoglob |
| Q88745 | mus musculu |
| Q928K6 | rattus norv |
| P04611 | human immun |
| P19552 | human immun |
| P08317 | gallus galli |
| P44037 | haemophilus |
| P26277 | erythrobact |
| Q91643 | streptococ |
| P06456 | avian papil |
| P38804 | saccharomyc |
| Q9vln5 | drosophila |
| Q75441 | homo sapien |
| Q02747 | homo sapien |
| P01820 | mus musculu |
| P20666 | oncorhynch |
| Q35282 | oncorhynch |
| P20667 | oncorhynch |
| Q37168 | oncorhynch |
| P11629 | oncorhynch |
| P20668 | oncorhynch |
| P25707 | oncorhynch |

| 107 | 108 | 109 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 | 129 | 130 | 131 | 132 | 133 | 134 | 135 | 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 | 144 | 145 | 146 | 147 | 148 | 149 | 150 | | | | |
|------------|------------|----------|------------|------------|------------|-----------|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----------|-----------|-----------|-------------|------------|------------|------------|------------|------------|-----------|-----------|------------|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|-----|-----|-----|-----|-----|
| 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | | |
| 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 |
| 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| NU3M_SALSA | NU3M_SALTR | YR4_PTV6 | B2MG_HUMAN | B2MG_PONPY | YPH5_CHRVI | RL7_HALEU | YG27_YEAST | EP4_RABIT | MIR1_HUMAN | RR12_SPINX | GAMD_DROME | YCD6_HUMAN | YMO6_SULSO | CCDP_MAIZE | OREX_MOUSE | OREX_MOUSE | R35A_AERPE | FLDC_SALTY | Y6C6_ECOLI | YL99_MYCTU | HBA_BRARE | HBA_CARAU | HBA_CARAU | ATRXA_STRCL | CAGA_STRCL | GSPG_XANCP | HV43_MOUSE | SLYA_ECOLI | SLYA_SALTY | HBB_CARAU | RL9_BACST | C554_RHOSH | PSAG_ARATH | PPI_STRCH | TCTP_HORVU | BMAE_ECOLI | LE23_METHH | CYSE_HELPJ | ECTA_MARNA | 3DHO_NECOR | YDES_ECOLI | RL10_SULSO | | | | | |

ALIGNMENTS

RESULT 1
EPA3_HUMAN
ID EPA3_HUMAN STANDARD; PRT; 238 AA.
AC P53797;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ephrin-A3 precursor (EPH-related receptor tyrosine kinase ligand 3)
DE (EPRK-3) (EHK1 ligand) (EHK1-L).
GN EFNA3 OR EPG3 OR LERK3 OR EFL-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95104019; PubMed=7838529;
RA Teese M., Lyman S.D., Srinivasan S., Fletcher F.A., Gayle R.B. III,
RA Cerretti D.P., Beckmann M.P.,
RT "Ligands for the receptor tyrosine kinases hek and elk: isolation of
RT cDNAs encoding a family of proteins."

SEQUENCE FROM N.A.
MEDLINE=95063919; PubMed=7973638;
RA Davis S., Gale N.W., Aldrich T.H., Maisonnier P.C., Lhotak V.,
RA Pawsan T., Goldfarb M., Yancopoulos G.D.,
RT "Ligands for EPH-related receptor tyrosine kinases that require
RT membrane attachment or clustering for activity."
RL Science 266:816-819(1994).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SKELETAL MUSCLE, SPLEEN,
CC THYMUS, PROSTATE, TESTIS, OVARY, SMALL INTESTINE, AND PERIPHERAL
CC BLOOD LEUKOCYTES.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL; U14187; AAC50078.1;
CC EMBL; L37360; AAA52368.1;
CC
CC MIM; 601381;
CC InterPro: IPR001799; Ephrin.
CC Pfam: PF00812; Ephrin.1.
CC PRINTS; PR01347; EPHRIN.
CC PRODOM; PD002533; Ephrin.
CC PROSITE; PS01299; EPHRIN.1.
CC KW Glycoprotein; GPI-anchor; Signal.
CC STGNAL 1 22
CC FT CHAIN 23 238
CC FT CARBOHYD 38 38
CC FT CARBOHYD 67 67
CC FT CARBOHYD 100 100
CC FT CONFLICT 71 74
CC SQ SEQUENCE 238 AA; 26350 MW; 8EFD6AE8FE33FD6A CRC64;

Query Match 1.6%; Score 8; DB 1; Length 238;
Best local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 AGPGPGG 86
DB 75 AGPGPGG 82

RESULT 2
SUT1_YEAST
ID SUT1_YEAST STANDARD; PRT; 299 AA.
AC P53032;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable sterol carrier.
DE SUT1 OR YGL162W OR G1828.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 28363 / F1100;
RX MEDLINE=96084960; PubMed=7489925;
RA Bourot S., Karst F.,
RT "Isolation and characterization of the Saccharomyces cerevisiae SUT1
RT gene involved in sterol uptake."
RL Gene 165:97-102(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96158061; PubMed=8585324;
RA James C.M., Indge K.J., Oliver S.G.,